Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/FI05/050007

International filing date: 17 January 2005 (17.01.2005)

Document type: Certified copy of priority document

Document details: Country/Office: FI
Number: 20040052
Filing date: 15 January 2004 (15.01.2004)

Date of receipt at the International Bureau: 07 March 2005 (07.03.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)

World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

PATENTTI- JA REKISTERIHALLITUS
NATIONAL BOARD OF PATENTS AND REGISTRATION

PCT/FI2005/050007

Helsinki 23.2.2005

E T U O I K E U S T O D I S T U S
P R I O R I T Y D O C U M E N T

Hakija Applicant	**Oy Jurilab Ltd** **Kuopio**
Patenttihakemus nro Patent application no	**20040052**
Tekemispäivä Filing date	**15.01.2004**
Kansainvälinen luokka International class	**C12Q**

Keksinnön nimitys
Title of invention

"Method for detecting the risk of acute myocardial infarction and coronary heart disease"
(Menetelmä akuutin sydäninfarktin ja sepelvaltimotaudin riskin havaitsemiseksi)

Täten todistetaan, että oheiset asiakirjat ovat tarkkoja jäljennöksiä Patentti- ja rekisterihallitukselle alkuaan annetuista selityksestä, patenttivaatimuksista ja tiivistelmästä.

This is to certify that the annexed documents are true copies of the description, claims and abstract, originally filed with the Finnish Patent Office.

Marketta Tehikoski
Apulaistarkastaja

Maksu 50 €
Fee 50 EUR

Maksu perustuu kauppa- ja teollisuusministeriön antamaan asetukseen 1142/2004 Patentti- ja rekisterihallituksen maksullisista suoritteista muutoksineen.

The fee is based on the Decree with amendments of the Ministry of Trade and Industry No. 1142/2004 concerning the chargeable services of the National Board of Patents and Registration of Finland.

L2

Method for detecting the risk of acute myocardial infarction and coronary heart disease

This invention relates to a method to detect genetic variation in a defensin gene for the diagnosis of a risk of, or predisposition to, acute myocardial infarction (AMI) and coronary heart disease (CHD) in a subject, a method for targeting treatment in a subject, and a method for selecting subjects for studies testing anticoronary agents, as well as a method for the treatment and prevention of CHD and AMI. The present invention also provides a method of identifying subject's susceptibility to or risk of developing AMI or CHD by detecting gene polymorphisms from a biological sample of the subject and obtaining information concerning the family and medical history, serum or plasma analytes and clinical findings of the subject. The invention also provides a multivariate model, a combination or algorithm of variables which best describes the probability of AMI and CHD. The invention also relates to a test kit and software for accomplishing the method. Moreover, the invention relates to a nucleic acid influencing the production of a novel variant defensin protein as well as a method for screening a subject to determine if said subject is a carrier of variant gene that encodes said variant or non-variant defensin protein.

FIELD OF THE INVENTION

The present invention is generally directed to a method for assessing the risk of CHD and AMI in an individual, such as a human. Specifically, the invention is directed to a method that utilises both genetic and phenotypic information as well as information obtained by questionnaires to construct a score that provides the probability of developing coronary heart disease. Furthermore, the invention provides a kit for carrying out the method. The kit can be used to set an etiology-based diagnosis of coronary heart disease and AMI for targeting of treatment and preventive interventions, such as dietary advice as well as stratification of the subject in clinical trials testing drugs and other interventions.

BACKGROUND OF THE INVENTION

Coronary heart disease (CHD) is the major cause of death in the developed world. The screening for conventional cardiovascular risk factors fails to identify more than 50% of the individuals who will present with acute coronary syndromes or AMI. Inflammation plays a role in both the development of atherosclerosis and the acute activation of the vascular wall with consequent local thrombosis and vasoconstriction. In many patients with unstable angina and AMI, systemic signs of inflammation are detectable. The use of systemic inflammatory markers, such as C-reactive protein as marker of disease activity and short- and long-term prognosis, seems to be of clinical value. Therefore, acute inflammatory reaction, detectable systematically, is a plausible risk factor for CHD and AMI.

As CHD is a polygenic disease, it is reasonable to assume that genetic variation in mechanisms important for the regulation of biochemical pathways that have a role in the development of atherosclerosis and CHD will be found to be associated with the pathogenesis and therapy of CHD.

One of the currently explored markers of inflammation is defensin. Defensins are a family of small cationic, antibiotic peptides that contain six cysteines in disulfide linkage. The peptides are abundant in phagocytes and small intestinal mucosa of humans and other mammals. They contribute to host defense against microbes and may participate in tissue inflammation and endocrine regulation during infection (Ganz and Lehrer 1995, Valore et al. 1998) and are a part of the innate immune system (Jia et al. 2001). There are two classes of defensin genes, α and β, that differ in their disulfide bond pairing, genomic organization, and in their tissue distributions. In addition to their broad spectrum antimicrobial properties, there is evidence that the β-defensins act as chemokines for immature dendritic cells and memory T cells, and thus may serve as an important bridge between the innate and adaptive immune systems (Jia et al. 2001, Hoover et al. 2001).

Defensins are normally sequestered in cytoplasmic granules with their primary site of action in phagolysosomes, although some peptide is released into the circulation during the course of infection or inflammation. Defensins have been found primarily in the intima of normal and atherosclerotic arteries, most prominently in association with intimal smooth

muscle cells by immunohistochemistry. Defensins are also found in the media near the external elastic lamina and in some periadventitial vessels. This indicates the presence of defensins in the walls of human coronary arteries. The deposition of defensins in vessels may contribute to the pathophysiological consequences of inflammation in addition to their role in host defense (Barnathan et al. 1997).

Characteristically, the antimicrobial activity of the β-defensin peptides is salt sensitive and their killing is markedly reduced as the ionic strength of the solutions increases (i.e., NaCl > 50 mM) (Schutte and McCray 2002).

The primary structure of each β-defensin gene product is characterized by small size, a six-cysteine motif, high cationic charge, and exquisite diversity beyond these features. The most characteristic feature of defensin proteins is their six-cysteine motif. Each β-defensin gene encodes a preproprotein that ranges in size from 59 to 80 amino acids with an average size of 65 amino acids. This gene product is then cleaved to create the mature peptide that ranges in size from 36 to 47 amino acids, with an average size of 45 amino acids (Schutte and McCray 2002) and molecular mass of 3-4 kD (Bensch at al. 1995).

At least 6 beta-defensins (HBD-1, HBD-2, HBD-3, HBD-4, HBD-5, HBD.6) have been characterized in humans. Human β-defensin-1 (HBD-1) was the first one to be characterized and isolated from the hemofiltrate of patients with end stage kidney disease undergoing dialysis (Lehmann et al. 2002). HBD-1 gene is expressed predominantly in urogenital epithelial organs such as kidney, urinary bladder, ureter and the female genital tract, with lesser expression in the pancreas, liver, and other epithelia. Within the kidney, in situ hybridization indicates that HBD-1 is produced in distal tubules, loops of Henle, and collecting ducts. Human urine contains 10-100 μg/L of HBD-1 (Zucht et al. 1998, Ganz 2001).

The human β-defensin-1 (HBD-1) gene covers approximately 8 kB on chromosome 8p23·1 (Dork and Stuhrmann 1998) and is comprised of two exons separated by an intron that is usually 1.5 kb, but can be as large as 16 kb. The processed transcript varies from 300 to 400 nucleotide (nt) in length with a 5' UTR 35 nt, an open reading frame of 200 nt, and a 3' UTR of 100 nt. The first exon includes the 5' UTR and encodes the leader domain of the

preproprotein; the second exon encodes the mature peptide with the six-cysteine domain (Schutte and McCray 2002).

Thus, inflammatory mechanisms are important participants in the pathophysiology of CHD. The identification of useful markers of inflammation and host resistance (like defensins), of new therapeutic targets to interfere with these mechanisms, and the evaluation of the efficacy of anti-inflammatory treatments will allow progress in our ability to prevent and manage CHD and combat its complications.

SUMMARY OF THE INVENTION

The object of this invention is to provide a method for screening a subject to assess if an individual is at risk to develop myocardial infarction or coronary heart disease, based on the genotype of a defensin gene and a method to target treatments and preventive therapies for CHD and AMI. The invention also provides methods for the treatment of CHD in a human or animal subject. A further object of the invention is to provide a method for the selection of experimental animals and human subjects for studies testing anticoronary and antihypertensive effects of drugs. Another object of the invention is a method for the selection of subjects for clinical trials testing anticoronary and antihypertensive drugs. A further object of the present invention is a method of identifying the risk of AMI and coronary heart disease by detecting gene polymorphisms from a biological sample of the subject. The information obtained from this method can be combined with other information concerning an individual, e.g. results from blood measurements, clinical examination and questionnaires. The genetic information includes data on mutations in genes associated with MI and/or coronary heart disease. The blood measurements include the determination of plasma or serum cholesterol and high-density lipoprotein cholesterol. The information to be collected by questionnaire includes information concerning gender, age, family and medical history such as the family history of CHD and diabetes. Clinical information collected by examination includes e.g. information concerning height, weight, hip and waist circumference, systolic and diastolic blood pressure, and heart rate.

More particularly, the invention provides a method for detecting genetic variation or polymorphism, i.e. a mutation, in a defensin gene comprising the steps of:

i) providing a biological sample taken from a subject to be tested,

ii) detecting the presence or absence of a variant genotype of the defensin gene in the biological sample, the presence of a variant defensin genotype indicating an increased risk of cardiovascular disease, such as CHD and AMI, in said subject.

Said defensin gene can be selected from the group consisting of: beta-defensin-1, beta-defensin-129, and alfa-defensin-5.

Moreover, the pattern of gene alleles can be further determined from the genes selected from the group consisting of:

a) alpha-$_{2B}$-adrenoceptor,
b) apolipoprotein B, and
c) beta-2-adrenergic receptor

in order to confirm the risk of cardiovascular disease in said subject.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS OF THE INVENTION

In a preferred embodiment the invention comprises the assessment of genetic variants in a defensin gene or the combination of information from a large number of variables (measurements) to predict the probability of AMI or CHD. The predictor information includes an assessment of genotypes and haplotypes in genomic DNA and optionally data obtainable by interviews, questionnaires, clinical examination and/or blood analyte measurements. This predictor information can be collected in any age. This method is also applicable to middle-aged persons.

The genetic, genotypic and phenotypic information used to predict AMI and CHD may relate to lipid, carbohydrate, amino acid and other nutrient (such as iron and folate) absorption, storage and metabolism, lipid transfer, oxidative and antioxidative metabolism, coagulation, fibrinolysis, platelet function, matrix proteins and degradation, blood pressure,

arterial contractility and constriction, other vasoregulation, renal function, central nervous system, properties of myocardium, glucose homeostasis, adiposity, arterial and myocardial cell necrosis, apoptosis, proliferation, migration and adhesion, inflammation (such as C-reactive protein), sympathetic tone such as adrenergic receptors or human host resistance against inflammation such as the defensins.

Numerous genotyping methods have been described in the art for analysing nucleic acids for the presence of specific sequence variations e.g. SNP's, insertions and deletions (for review see Syvänen 2001 and Nedelcheva Kristensen et al. 2001). In these methods a sample containing nucleic acid (e.g. blood, tissue biopsy or buccal cells) is obtained from the patient and the sequence variations of interest are identified and assessed from the nucleic acids.

Allelic variants in genes can be discriminated by enzymatic methods (with the aid of restriction endonucleases, DNA polymerases, ligases etc.), by electrophoretic methods (e.g. single strand conformation polymorphism (SSCP), heteroduplex analysis, fragment analysis and DNA sequencing), by solid-phase assays (dot blots, microarrays, microparticles, microtiter plates etc.) and by physical methods (e.g. hybridisation analysis, mass spectrometry and denaturing high performance liquid chromatography (DHPLC)). In most of the genotyping assays different polymerase chain reaction (PCR) applications are used both to increase the signal to noise ratio as well as spare sample nucleic acid before allele discrimination. Detectable labels (fluorochromes, radioactive labels, biotin, modified nucleotides, haptens etc) can be used to enhance visualization of allelic variants.

This invention is based on the principle that one or a small number of genotypings are performed, and the mutations to be typed are selected on the basis of their ability to predict AMI and/or CHD. For this reason any method to genotype mutations in a genomic DNA sample can be used. If non-parallel methods such as real-time PCR are used, the typings are done in a row. The PCR reactions may be multiplexed or carried out separately in a row or in parallel aliquots.

The score that predicts the probability of MI or CHD may be calculated using a multivariate failure time model or a logistic regression equation as follows:

Probability of coronary heart disease = $[1 + e^{(-(-a + \Sigma(bi*Xi))}]^{-1}$, wherein e is Napier's constant, X_i are variables related to the cardiovascular disease, b_i are coefficients of these variables in the logistic function, and a is the constant term in the logistic function. The model may additionally include any interaction (product) or terms of any variables X_i, e.g. b_iX_i. An algorithm is developed for combining the information to yield a simple prediction of MI as percentage of risk in 10 years. Alternative statistical models are a failure-time models such as the Cox's proportional hazards' model and neural networking models.

Thus, the detection method of the invention may further comprise a step of combining information concerning age, gender, the family history of hypertension, diabetes and hypercholesterolemia, and the medical history concerning cardiovascular diseases or diabetes of the subject with the results obtained from step ii) of the method (see claim 1) for confirming the indication obtained from the detection step. Said information may also concern hypercholesterolemia in the family, smoking status, CHD in the family, history of cardiovascular disease, obesity in the family, and waist-to-hip circumference ratio (cm/cm)

The detection method of the invention may also further comprise a step determining blood, serum or plasma cholesterol, HDL cholesterol, LDL cholesterol, triglyceride, apolipoprotein B and AI, fibrinogen, ferritin, transferrin receptor, C-reactive protein, serum or plasma insulin concentration.

The results from the furher steps of the method as described above render possible a step of calculating the probability of a cardiovascular disease using a logistic regression equation as follows:

Probability of a cardiovascular disease = $[1 + e^{(-(-a + \Sigma(bi*Xi))}]^{-1}$, where e is Napier's constant, X_i are variables related to the cardiovascular disease, b_i are coefficients of these variables in the logistic function, and a is the constant term in the logistic function, and wherein a and b_i are preferably determined in the population in which the method is to be used, and Xi are preferably selected among the variables that have been measured in the population in which the method is to be used. Preferable values for b_i are between –20 and 20; and for i between 0 (none) and 100,000. X_i are binary variables that can have values or are coded as 0 (zero) or 1 (one).

The method can be used in the prediction and early diagnosis of AMI in adult persons, stratification and selection of subjects in clinical trials, stratification and selection of persons for intensified preventive and curative interventions. The aim is to reduce the cost of clinical drug trials and health care.

The test can be applied to test the risk of developing an AMI in both

1) healthy persons, as a screening or predisposition test and
2) high-risk persons (who have e.g. family history of CHD or elevated serum cholesterol or hypertension or diabetes or any combination of these).

As inflammation is a cause of AMI and other forms of CHD, anti-inflammatory agents can plausibly be used in the prevention and treatment of AMI and chronic CHD. Persons who have a compromised host resistance to inflammation, due to e.g. reduced expression or production of human defensin proteins, will thus benefit from defensin enhancing medications, diets and other therapies. More generally, all people might benefit from the enhancement of the defensin system through a reduction of their AMI and CHD risk and consequent increase in longevity. Especially persons whose defensin levels are lowered or who have mutations in the genes encoding human defensins will benefit from such a treatment. Other groups or persons which will get increased benefit from defensin enhancing treatments are persons who already have CHD. Clinical trial testing the effect of defensing enhancement on defensin expression, body defensin levels, the progression of atherosclerosis and the incidence of AMI and other coronary events can be carried out with compounds enhancing body defensin levels and methods to measure said compounds. A method for treating a human or animal suffering from CHD or AMI by enhancing defensin availability, production or concentration in the human subject or animal may comprise an administration of a chemical entity such as a medication, a vaccination, a nutrient in natural or functional food or foodstuff, other behavioural intervention or gene therapy such as gene transfer.

As defensins are necessary in protecting against CHD and AMI, medications, dietary and other treatments that reduce human defensin levels or activity will cause adverse reactions in those persons. The likelihood of adverse reactions is the greatest in persons who already have lowered defensin levels or activities.

Transgenic animal models with mutant defensin genes and defensin gene knock-out animal models can be used to study the effect and role of defensins in the causation and progression of AMI, CHD and other diseases and conditions. RNA interference of defensin genes may be used to for the same purposes. As these model animals have increased susceptibility to CHD, they can also be used to study the efficacy and adverse reactions of any medication, nutrient or other compound in the treatment or prevention of AMI and CHD.

More particularly, the invention is directed to a method for detecting genetic variation or polymorphism, i.e. a mutation, in a defensin gene comprising the steps of:

i) providing a biological sample taken from a subject to be tested,

ii) detecting the presence or absence of a variant genotype of the defensin gene in the biological sample, the presence of a variant defensin genotype indicating an increased risk of cardiovascular disease in said subject.

Preferably, genetic variation is further determined from the genes selected from the group consisting of:

a) alpha-$_{2B}$-adrenoceptor,
b) apolipoprotein B, and
c) beta-2-adrenergic receptor

wherein the presence of a variant genotype in said genes indicates an increased risk of cardiovascular disease, such as myocardial infarction (AMI) or coronary heart disease (CHD), in said subject.

The method may further comprise a step of selecting a subject with a variant defensin gene sequence reducing the expression, production or levels of defensin protein for clinical drug trials testing the anticoronary and myocardial ischaemia preventing effects of compounds.

Preferably, said variant genotype of the defensin gene is a homo- or heterozygote form of the mutation.

The detection step of the method can be a DNA-assay. Such detection step can also be carried out using a gene or DNA chip, microarray, strip, panel or similar combination of more than one genes, mutations or RNA expressions to be assayed. Moreover, one of the preferable embodiments of the invention is the determination of the allelic pattern by polymerase chain reaction. The detection step of the method can also be based on a capturing probe, which specifically binds to a variant defensin nucleic acid.

The biological sample for the method can be, e.g., a blood sample or buccal swab sample. From said sample genomic DNA is isolated.

The subject to be tested is preferably a mammal, more preferably a primate, and most preferably a human.

The method of the invention can be used used for determining whether a subject will benefit from treatment with a drug, nutrient or other therapy enhancing the defensin production, levels or activity or inhibiting defensin catabolism or elimination in the subject. Moreover, the method can be preferably used for determining whether a subject will be at increased risk of adverse effects or reactions if defensin antagonists are administered to a subject.

The method of the invention is preferably directed to the detection of the variants of the following genes: human beta-defensin-1 (e.g. 3`UTR +5A→G variant), human beta-defensin-129 (e.g. 5'UTR –27T→C variant and/or IVS1 –13_12insCTC), human alfa-defensin-5 (e.g. IVS1 +198C→T variant and/or IVS1 +243G→C variant), beta-2-adrenergic receptor (e.g. Gly16Arg variant and/or Glu27Gln variant) and alpha-2B-adrenergic receptor (e.g.insertion/deletion variant as defined in the Experimental Section), and apolipoprotein B gene (e.g. Thr98Ile variant). Thus, the listed gene variants are shown herein to predict CHD and/or AMI. However, a person skilled in the art may find by routine work new functional mutations in said genes. Such variants are deemed to be within the scope of those skilled in the art from the teachings herein.

The present invention also provides a method for targeting the treatment of CHD, such as angina pectoris or other form of CHD, and AMI in a subject with CHD by determining the pattern of alleles encoding a variant defensin, i.e. by determining if

said subject's genotype of the defensin is of the variant type, comprising the steps presented in claim 1, and treating a subject of the variant genotype with a drug affecting defensin production or metabolism of the subject.

Another embodiment of the invention is a method for treating a human or animal suffering from CHD or AMI, said method comprising a therapy enhancing defensin availability, production or concentration of the human subject or animal, such as a mammal. Such method can be, e.g., for treating vascular complications of CHD and AMI, wherein said method may comprise a step of enhancing defensin availability, production or concentration in the circulation of a human subject or animal. The treatment may be, e.g., a dietary treatment, a vaccination, gene therapy or gene transfer. Said gene therapy may comprise a transfer of a non-variant defensin gene, such as beta-defensin-1, or fragment or derivative thereof.

The present invention further provides a kit for detecting genetic variation or polymorphism, i.e. a mutation, in a defensin gene for the determination of a risk of acute myocardial infarction, AMI, and coronary heart disease, CHD, in a subject, comprising means for defensin gene allele detection, and optionally software and/or instructions to interpret the results of the determination. The kit may also provide means for the detection of the variants of the genes selected from the group consisting of:

a) alpha$_{-2B}$-adrenoceptor,
b) apolipoprotein B, and
c) beta-2-adrenergic receptor

Preferably, the detected variants are the ones as described above and in the Experimental Section.

The kit can be based on a capturing nucleic acid probe specifically binding to the variant genotype as defined in the invention, and/or on a DNA chip, microarray, DNA strip, DNA panel or real-time PCR based tests. The kit may also comprise a questionnaire for obtaining patient information concerning age, gender, height,

weight, the family history of hypertension and hypercholesterolemia, the medical history concerning cardiovascular diseases.

The publications and other materials used herein to illuminate the background of the invention, and in particular, to provide additional details with respect to its practice, are incorporated herein by reference.

The invention will be described in more detail in the Experimental Section.

EXPERIMENTAL SECTION

Determining individual genotypes

For the identification of the specific gene mentioned in the experimental section we have used Locus Link ID numbers (http://www.ncbi.nlm.nih.gov/LocusLink/). For the identification of the specific known SNPs mentioned in the experimental section we have used rs-numbers from the NCBI SNP database (http://www.ncbi.nlm.nih.gov/SNP/)

The method according to the invention for the determination of the allelic pattern of the DNA variation in question can be carried out with polymerase chain reaction (PCR) in combination with an allele specific primer extension method (SNaPshot, Applied Biosystems) followed by capillary electrophoresis with ABI Prism 3100 Genetic Analyzer (Applied Biosystems).

In a snapshot reaction the genomic DNA region containing the variation in question is amplified with PCR. The amplified PCR product is purified and used as a template in the snapshot reaction. For the snapshot reaction an extension primer is designed so that the 3' end of the primer is immediately adjacent to the polymorphic site of interest. In the snapshot reaction the extension primer hybridizes to its complementary template in the presence of fluorescent labeled dideoxy-NTPs ([F]ddNTPs) and DNA polymerase. The polymerase extends the primer by only one nucleotide, adding a single [F]ddNTP to its 3' end. Because each of the four [F]ddNTPs are labeled with different fluorescent dyes the genotypes can be discriminated.

If multiple SNPs are to be determined in the same reaction, the extension primers need to be designed so that they differ from each other significantly in length (4-6 nucleotides). The length of a primer can be modified by the addition of a variable, but a known number of non-homologous nucleotides (dT, dA, dC or cGATC) to the 5' end of the extension primers. Due to the difference in the length of the extension primers the snapshot products can be detected in the capillary electrophoresis according to the size of the product. To perform SnaPshot genotyping under standard conditions, refer to the user manual (ABI Prism SnaPshot Multiplex kit, Protocol, Applied Biosystems).

Polymerase chain reaction (PCR)

The genomic DNA regions containing the mutations in question can be amplified with PCR either in separate reactions or all in one single reaction mix (i.e. multiplex PCR). The PCR amplification was conducted in a 30 μl volume: the reaction mixture contained 40 ng human genomic DNA (extracted from peripheral blood), 1X PCR Buffer (QIAGEN), 200 μM of each of the nucleotides (dATP, dCTP, dGTP, dTTP, Finnzymes), 0.75 μM of DEFB1 PCR primers, 0.5 μM of DEFB129 and DEFA5 PCR primers and 0.25 μM of ADRB2 PCR primers and 2.5 units of Hot Start Taq DNA polymerase (QIAGEN). First the reaction was hold 5 minutes at 96°C, then the following three steps were repeated for 35 cycles: 30 seconds at 94°C, 1 minute at 57°C, 1 minute at 72°C, after which the reaction was kept at 72°C for an additional 5 minutes and then hold 1 minute 10°C and stored at 4°C in a PTC-220 DNA Engine Dyad PCR machine (MJ Research).

For the APOB the PCR amplification was conducted in a 20 μl volume: the reaction mixture contained 40 ng human genomic DNA (extracted from peripheral blood), 1X PCR Buffer (QIAGEN), 200 μM of each of the nucleotides (dATP, dCTP, dGTP, dTTP, Finnzymes), 10 pmol of APOB PCR primers and 2.0 units of Hot Start Taq DNA polymerase (QIAGEN,). First the reaction was hold 7 minutes at 94°C, then the following three steps were repeated for 35 cycles: 45 seconds at 94°C, 45 seconds at 54°C, 1 minute at 72°C, after which the reaction was kept at 72°C for an additional 5 minutes and then hold 1 minute 10°C and stored at 4°C in a PTC-220 DNA Engine Dyad PCR machine (MJ Research).

For the DEFB129 IVS1 –12_13insCTC the amplification was conducted in a 40 μl volume: the reaction mixture contained 60 ng human genomic DNA (extracted from peripheral blood), 1X PCR Buffer (QIAGEN), 200 μM of each of the nucleotides (dATP, dCTP, dGTP, dTTP, Finnzymes), 20 pmol of DEFB129 IVS1 –12_13insCTC PCR primers and 3.0 units of Hot Start Taq DNA polymerase (QIAGEN,). First the reaction was hold 7 minutes at 96°C, then the following three steps were repeated for 35 cycles: 45 seconds at 94°C, 45 seconds at 57°C, 1 minute at 72°C, after which the reaction was kept at 72°C for an additional 5 minutes and then hold 1 minute 10°C and stored at 4°C in a PTC-220 DNA Engine Dyad PCR machine (MJ Research).

The nucleotide sequence of the PCR primer pair for the amplification of the human DEFB1 gene (defensin beta 1, Locus link ID: 1672) 3`UTR +5A>G mutation (rs1047031) was as follows: 5'- CAT AAT TTC AGC CCG ATG TG -3' (SEQ ID NO:1) and 5'- CAC CCT AAC CCC CTA CTT CT-3' (SEQ ID NO:2).

The nucleotide sequence of the PCR primer pair for the amplification of the human DEFB129 gene (defensin beta 129, Locus link ID: 140881) 5'UTR–27T>C (rs2298148) was as follows: 5'- GGG CTT GCT CTT TCT TTC -3' (SEQ ID NO:3) and 5'- TCC TTG GTT CCT CTC ATC -3' (SEQ ID NO:4).

The nucleotide sequence of the PCR primer pair for the amplification of the human ADRB2 gene (Beta-2-adrenergic receptor, Locus link ID: 154) Gly16Arg (rs1042713) and Glu27Gln (rs1042714) mutations was as follows: 5'- CTG AGT GTG CAG GAC GAG –3' and (SEQ ID NO:5) 5'- CAC ATT GCC AAA CAC GAT –3' (SEQ ID NO:6).

The nucleotide sequence of the PCR primer pair for the amplification of the human DEFA5 gene (defensin alpha 5, Locus link ID: 1670) IVS1 +198C>T (in the following sequence, [SEQ ID NO:7, SEQ ID NO:8], the DEFA5 IVS +198C>T substitution is located at the position 553) and the IVS1 +243G>C variants (in the following sequence, [SEQ ID NO:7, SEQ ID NO:8], the DEFA5 IVS +243G>C substitution is located at the position 598) was as follows: 5'- AGA AAG AGG AGC ATC AAA G -3' (SEQ ID NO:9) and 5'- TCA AGC CTA TTA GCC TAC A-3' (SEQ ID NO:10).

The nucleotide sequence of the PCR primer pair for the amplification of the human APOB gene (apolipoprotein B, Locus link ID: 338) Thr98Ile mutation (also known as Thr71Ile mutation, rs1367117) was as follow: 5'- GAC AAC CTC AAT GCT CTG CT -3' (SEQ ID NO:11) and 5'- TGA CTT ACC TGG ACA TGG CT -3' (SEQ ID NO:12).

The nucleotide sequence of the PCR primer pair for the amplification of the human DEFB129 gene (defensin beta 129, Locus link ID:140881) IVS1-12_13insertionCTC variant (in the following sequence, SEQ ID NO:32, SEQ ID NO:33 the insertion is in position 444-446) was as follow: 5'- GGC TAC TGA GTT TGG TGA -3' (SEQ ID NO:34) and 5'- GTG TTT ATT GAA TGA CTG ATG -3' (SEQ ID NO:35).

The PCR products were purified with SAP (Shrimp Alkaline Phosphatase, USB) and *Exo*I (Exonuclease I, New England Biolabs) treatment. This was done to avoid the participation of the unincorporated dNTPs and primers from the PCR reaction to the subsequent primer-extension reaction. More specifically, 2.5µl of SAP (1 unit/µl, USB), 0.25 µl of *Exo*I (20 units/µl, New England Biolabs), 1.0 µl of 10 X *ExoI* buffer (New England Biolabs) and 6.25 µl H_2O were added to 5 µl of the PCR product. Reaction was mixed and incubated at 37°C for 1 hour. After that the reaction was kept at 75°C for 15 minutes to inactivate the enzymes and stored at 4°C.

In the subsequent primer extension reaction (SNaPshot reaction) 1.5 µl of SNaPshot Multiplex Ready Reaction Mix (Applied Biosystems), 3 µl of purified PCR products, 1 µl of pooled extension primers (1 µM each) and 4.5 µl buffer (1X AmpliTaq Gold buffer 2mM $MgCl_2$, Applied Biosystems) are mixed in a tube. The reaction is incubated at 96°C for 5 seconds and then subject to 35 cycles of 95°C for 10 s, 50°C for 5 s and 60°C for 30 s in a PTC-220 DNA Engine Dyad PCR machine (MJ Research).

The nucleotide sequence of the extension primer for the genotyping of human DEFA5 IVS1 +198C>T mutation in a SNaPShot reaction was: 5'- TTT TTT TTT TTT TTT CTT TTT TCT AAG ACT TTC AG -3' (SEQ ID NO:13).

The nucleotide sequence of the extension primer for the genotyping of human DEFA5 IVS1 +243G>C mutation in a SNaPShot reaction was: 5'- TTT TTT TTT TTT TTT TTT TGC TAC TTT TAA GAT AGA AAG A -3' (SEQ ID NO:14).

The nucleotide sequence of the extension primer for the genotyping of the human DEFB1 3`UTR +5A>G mutation in a SNaPShot reaction was: 5'- TTT TTT TTT TTT TTT TTT TTT TTT TTT AGT GCT GCA AGT GAG CTG –3' (SEQ ID NO:15).

The nucleotide sequence of the extension primer for the genotyping of human DEFB129 IVS1 –27T>C mutation in a SNaPshot reaction was: 5'- TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT CCA GAG AGG AAG CCT TG-3' (SEQ ID NO:16).

The nucleotide sequence of the extension primer for the genotyping of human ADRB2 Gly16Arg mutation in a SNaPShot reaction was: 5'- T TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTC TTG CTG GCA CCC AAT –3' (SEQ ID NO:17).

The nucleotide sequence of the extension primer for the genotyping of human ADRB2 Glu27Gln mutation in a SNaPShot reaction was: 5'- TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TAC CAC GAC GTC ACG CAG –3' (SEQ ID NO:18).

The nucleotide sequence of the extension primer for the genotyping of human APOB Thr98Ile (Thr71Ile, rs1367117) mutation in a SNaPShot reaction was: 5'- TTT TTT TTT TTT TGA AGA CCA GCC AGT GCA –3' (SEQ ID NO:19).

The nucleotide sequence of the extension primer for the genotyping of human DEFB129 gene (defensin beta 129) IVS1-12_13insertionCTC variant in a SNaPshot reaction was: 5'- TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT GCT CAA TGG CTT TCT CT -3' (SEQ ID NO:56). In the snapshot reaction the deletion CTC allele is detected as nucleotide T whereas the presence of the insertion CTC allele is detected as nucleotide C.

After the primer extension reaction (snapshot reaction) 1 unit of SAP (USB) was added to the reaction mix and the reaction was incubated at 37°C for 1 hour. The enzyme was inactivated by incubating the reaction mix at 75°C for 15 minutes and placed at 4°C. The post-extension treatment was done to prevent the unincorporated fluorescent ddNTPs obscuring the primer extension products (SNaPshot products) during electrophoresis with ABI Prism 3100 Genetic Analyzer.

DNA fragment analysis of ADRA2B insertion/deletion polymorphism

The insertion/deletion polymorphism of ADRA2B gene concerns an insertion or a deletion of three glutamic acids (Glu) in the region of 12 Glu amino acids in the codons 298-309. Thus depending on the allele, there is either 9 Glu (deletion, variant form) (SEQ ID NO:20) or 12 Glu (insertion) (SEQ ID NO:22) at the ADRA2B locus. Depending on whether the amplified allele had an insertion or a deletion in the studied locus, the size of the PCR product was 91 bp (insertion allele) or 82 bp (deletion allele). For homotzygotes (insertion/insertion or deletion/deletion) only one size of a fragment was detected either 91 bp or 82 bp, respectively. For heterotzygotes both of the above mentioned fragments were detected.

The PCR amplification was conducted in a 20 µl volume: the reaction mixture contained 40 ng human genomic DNA (extracted from peripheral blood), 1X PCR Buffer (QIAGEN), 200 µM of each of the nucleotides (dATP, dCTP, dGTP, dTTP), 10 pmol of ADRA2B PCR primers and 2.0 units of Hot Start Taq DNA polymerase (QIAGEN). First the reaction was hold 7 minutes at 95°C, then the following three steps were repeated for 35 cycles: 45 seconds at 94°C, 45 seconds at 54°C, 1 minute at 72°C, after which the reaction was kept at 72°C for an additional 5 minutes and then hold 1 minute 10°C and stored at 4°C in a PTC-220 DNA Engine Dyad PCR machine (MJ Research).

The PCR primer pair for the amplification of the ADRA2B gene (alpha-2B-adrenergic receptor, Locus link ID: 151) insertion/deletion polymorphism was as follows 5'– GGG TGT TTG TGG GGC ATC TC –3' (SEQ ID NO:24) and 5'- TGG CAC TGC CTG GGG TTC A -3' (SEQ ID NO:25). A fluorescent label has been added to the 5' end of one of the above mentioned PCR primers. Therefore, the PCR fragment is detectable in the capillary electrophoresis conducted with ABI Prism 3100 Genetic Analyzer (Applied Biosystems).

Capillary electrophoresis with ABI Prism 3100 Genetic Analyzer

Aliquots of 1 µl of pooled SNaPshot products, 0.5 µl ADRA2B PCR product, 9.25 µl of Hi-Di formamide (Applied Biosystems) and 0.25 µl GeneScan-120 LIZ size standard (Applied Biosystems) were combined in a 96-well 3100 optical microamp plate (Applied Biosystems). The reactions were denatured by placing them at 95°C for 5 minutes and then

loaded onto a ABI Prism 3100 Genetic Analyzer (Applied Biosystems). Electrophoresis data was processed and the genotypes were visualized by using the GenoTyper Analysis Program version 3.7 (Applied Biosystems).

Identification of new mutations in human beta-defensin genes

We used the hierarchial phenotype-targeted sequencing method (see WO 02/074230) to find new mutations in the beta-defensin-1 gene. As defensins are known to act to protect against infections, it was hypothesised that subjects with frequent infections would have lowered and subjects with infrequent infections would have high or normal body defensin levels and activities. Forty-eight Kuopio Ischaemic Heart Disease Risk Factor Study (KIHD) examinees with the largest number of respiratory and urinary infections in the previous five years and 48 gender- and age-matched subjects with neither respiratory nor urinary infections in the previous five years were selected for sequencing. We sequenced five different Defensin Alpha genes (DEFA1, DEFA3, DEFA4, DEFA5 and DEFA6) and six different Defensin Beta genes (DEFB1, DEFB103, DEFB4, DEFB118, DEFB126 and DEFB129).

In sequencing we found five mutations in DEFA5 gene (DEFA5 IVS1 +198C>T, DEFA5 IVS1 +243G>C, DEFA5 Arg71Cys [rs7839771], DEFA5 3`UTR +109A>G and DEFA5 3`UTR +168C>T).

In DEFB1 gene we found five mutations (DEFB1 5`UTR-52G>A [rs1799946], DEFB1 5`UTR–44C>G [rs1800972], DEFB1 5`UTR–20A>G [rs11362], DEFB1 IVS1+19T>A [rs2293958] and DEFB1 3`UTR+5A>G [rs1047031]).

In DEFB2 gene we found three mutations (DEFB2 5`UTR–108T>C [rs2740086], DEFB2 T>C Pro29Pro [rs2740090] and DEFB2 3`UTR+164G>A [rs2737531]).

From DEFB118 gene we found one mutation (DEFB118 T>C Cys34Arg).

In DEFB126 gene we found two mutations (DEFB126 deletion CAAA163_166 frameshift and DEFB126 deletion CC317_318 frameshift).

In DEFB129 gene we found five mutations (DEFB129 5`UTR–41G>A [rs2298149], DEFB129 5`UTR–27T>C [rs2298148], DEFB129 IVS1–68C>T [rs6074833], DEFB129 IVS1 –13_12insertionCTC and DEFB129 A201G synonymous to Leu67Leu).

Of the above mentioned Defensin Alpha and Defensin Beta gene variants the following 9 (nine) have not been reported previously: DEFA5 IVS1+198 C>T, DEFA5 IVS1+243 G>C, DEFA5 3'UTR+109 A>G, DEFA5 3'UTR+168 C>T, DEFB129 IVS1-12 insertion deletion CTC, DEFB129 A>G leu67leu (CTG67CTA), DEFB118 T>C Cys34Arg (TG34CGC), DEFB126 exon 2 deletion c.163_ 166delCAAA and DEFB126 exon 2 deletion c.317_318delCC.

The nucleotide sequence of the PCR primer pair for the amplification of the human DEFA5 gene (defensin alpha 5) IVS1+198 C>T variant (in the following sequence, [SEQ ID NO:7, SEQ ID NO:8)] the substitution is located at the position 553) was as follow: 5'- AGA AAG AGG AGC ATC AAA G -3' (SEQ ID NO:9) and 5'- TCA AGC CTA TTA GCC TAC A -3' (SEQ ID NO:10). The sequencing primer was: 5' – TCA GGT CTT CTC CCA GCA (SEQ ID NO:26)

The nucleotide sequence of the PCR primer pair for the amplification of the human DEFA5 gene (defensin alpha 5, Locus link ID:1670) IVS1+243 G>C variant (in the following sequence, [SEQ ID NO:7, SEQ ID NO:8)] the substitution is located at the position 598) was as follow: 5'- AGA AAG AGG AGC ATC AAA G -3' (SEQ ID NO:9) and 5'- TCA AGC CTA TTA GCC TAC A -3' (SEQ ID NO:10). The sequencing primer was: 5' – TCA GGT CTT CTC CCA GCA (SEQ ID NO:26)

The nucleotide sequence of the PCR primer pair for the amplification of the human DEFA5 gene (defensin alpha 5, Locus link ID:1670) 3'UTR+109 A>G variant (in the following sequence, [SEQ ID NO:27, SEQ ID NO:28] the substitution is located in position 515) was as follow: 5'- GGA TGA AGC AGA ATG AAG A -3' (SEQ ID NO:29) and 5'- AAA GGA ACC ATA CAA ACC A -3' (SEQ ID NO:30). The sequencing primer was: 5' – GTT AGT CTG GCT GTG CTT – 3' (SEQ ID NO:31).

The nucleotide sequence of the PCR primer pair for the amplification of the human DEFA5 gene (defensin alpha 5, Locus link ID:1670) 3'UTR+168 C>T variant (in the

following sequence, [SEQ ID NO:27, SEQ ID NO:28] the substitution is located in position 574) was as follow: 5'- GGA TGA AGC AGA ATG AAG A -3' (SEQ ID NO:29) and 5'- AAA GGA ACC ATA CAA ACC A -3' (SEQ ID NO:30). The sequencing primer was: 5' – GTT AGT CTG GCT GTG CTT – 3' (SEQ ID NO:31).

The nucleotide sequence of the PCR primer pair for the amplification of the human DEFB129 gene (defensin beta 129, Locus link ID:140881) IVS1-12_13insertionCTC variant (in the following sequence, SEQ ID NO:32, the insertion is in position 444-446) (SEQ ID NO:33) was as follow: 5'- GGC TAC TGA GTT TGG TGA -3' (SEQ ID NO:34) and 5'- GTG TTT ATT GAA TGA CTG ATG -3' (SEQ ID NO:35). The sequencing primer was: 5' – CAA GGA AGG CAG ACT AAA – 3' (SEQ ID NO:36).

The nucleotide sequence of the PCR primer pair for the amplification of the human DEFB129 gene (defensin beta 129, Locus link ID:140881) leu67leu (CTA67CTG), A>G variant (SEQ ID NO:37) (SEQ ID NO:39) was as follow: 5'- GGC TAC TGA GTT TGG TGA -3' (SEQ ID NO:34) and 5'- GTG TTT ATT GAA TGA CTG ATG -3' (SEQ ID NO:35). The sequencing primer was: 5' – CAA GGA AGG CAG ACT AAA – 3' (SEQ ID NO:36).

The nucleotide sequence of the PCR primer pair for the amplification of the human DEFB118 gene (defensin beta 118, Locus link ID:117285) Cys34Arg (TGC34CGC), T>C mutation (SEQ ID NO:41) (SEQ ID NO:43) was as follow: 5'- AGG TTG AGT ATT TGC CAG AC -3' (SEQ ID NO:45) and 5'- AGG ACA GGG GTG AGT GAT A -3' (SEQ ID NO:46). The sequencing primer was: 5' – AGG TTG AGT ATT TGC CAG AC – 3' (SEQ ID NO:45).

The nucleotide sequence of the PCR primer pair for the amplification of the human DEFB126 (defensin beta 126, Locus link ID:81623) exon 2 deletion c.163_166delCAAA (SEQ ID NO:47) (SEQ ID NO:49) was as follow: 5'- AAT GGT GAG AAA GAT GAC AG -3' (SEQ ID NO:51) and 5'- GTT GAA TGG AGG GAA AGT -3' (SEQ ID NO:52). The sequencing primer was: 5' – GTA GGT ATT TAT GAT TAG – 3' (SEQ ID NO:53). This mutation leads to a change in protein amino acid structure of the DEFB126 gene from the amino acid codon 55 and finally to a premature STOP codon in amino acid position 82 (SEQ ID NO:47).

The nucleotide sequence of the PCR primer pair for the amplification of the human DEFB126 gene (defensin beta 126, Locus link ID:81623) exon 2 deletion c.317_318delCC (SEQ ID NO:54) (SEQ ID NO:49) was as follow: 5'- AAT GGT GAG AAA GAT GAC AG -3' (SEQ ID NO:51) and 5'- GTT GAA TGG AGG GAA AGT -3' (SEQ ID NO:52). The sequencing primer was: 5' – GTA GGT ATT TAT GAT TAG – 3' (SEQ ID NO:53). This mutation also leads to an altered amino acid structure of the DEFB126 gene from the amino acid codon 106 (SEQ ID NO:54).

Testing the Risk of AMI and CHD

Risk factors for MI and coronary heart disease were studied in the KIHD cohort. Briefly, the "Kuopio Ischaemic Heart Disease Risk Factor Study" (KIHD) is a prospective population study in men in Eastern Finland (Salonen 1988, Tuomainen et al. 1999). The study protocol for KIHD was approved by the Research Ethics Committee of the University of Kuopio. The study sample comprised men from Eastern Finland aged 42, 48, 54 or 60 years. A total of 2682 men were examined during 1984-89. All participants gave a written informed consent. The follow-up of coronary events was to the end of 2001, providing an average follow-up time of 14.4 years. Genotypings were carried out for approximately 1600 men, resulting to over 23,000 person-years of follow-up.

Data on CHD and AMI during the follow-up were obtained by computer record linkage to the national computerized hospital discharge registry. Diagnostic information was collected from the hospitals and all heart attacks events were classified according to rigid predefined criteria. The diagnostic classification of acute coronary events was based on symptoms, electrocardiographic findings, cardiac enzyme elevations, autopsy findings and the history of CHD. Each suspected coronary event (ICD-9 codes 410-414 and ICD-10 codes I20-I25) was classified into 1) a definite acute myocardial infarction (AMI), 2) a probable AMI, 3) a typical acute chest pain episode of more than 20 minutes indicating CHD, 4) an ischemic cardiac arrest with successful resuscitation, 5) no acute coronary event or 6) an unclassifiable fatal case. The categories 1) to 3) were combined for the present analysis to denote MI. Of 1548 male subjects with complete data, used in the analysis, 256 men developed an AMI during the follow-up.

Hypertension was defined as either systolic blood pressure (BP) ≥165 mmHg or diastolic BP ≥95 mmHg or antihypertensive treatment. Both blood pressures were measured in the morning by a nurse with a random-zero mercury sphygmomanometer. The measuring protocol included three measurements in supine, one in standing and two in sitting position with 5-minutes intervals. The mean of all six measurements were used as systolic and diastolic blood pressures. Family history of CHD was defined positive if either the subject's mother, father or a sibling had a history of AMI or angina pectoris. Family histories of cerebrovascular stroke and diabetes were defined similarly. Adulthood socioeconomical status (SES) is an index comprised of measures of education, occupation, income and material living conditions. The scale is inverse, low score corresponding to high SES. These data have been collected by a self administered questionnaire. Serum ferritin was assessed with a commercial double antibody radioimmunoassay (Amersham International, Amersham, UK). Lipoproteins, including high density lipoprotein (HDL) and low density lipoprotein (LDL), were separated from fresh serum samples by ultracentrifugation followed by direct very low density lipoprotein (VLDL) removal and LDL precipitation. Cholesterol concentration was then determined enzymically. Serum C-reactive protein was measured by a commercial high-sensitive immunometric assay (Immulite High Sensitivity CR Assay, DPC, Los Angeles). Genotyping of the paraoxonase 1 and HFE (HLA-H) mutations have been de4scribed elsewhere (Salonen et al. 1999, Tuomainen et al 1999).

In the beta-defensin 1 gene, 3'UTR+5, of the 1548 men genotyped, 165 were AA homozygotes, 690 heterozygotes and 693 GG homozygotes. Of the GG homozygotes, 19.0% (132 men) developed their first AMI during the follow-up, as compared with 14.5% (124 men) of the other men (odds ratio 1.39, 95% CI 1.06 to 1.82, p=0.017). In a multivariate logistic model controlling for the strongest other covariates, the respective adjusted odds ratio was 1.35 (95% CI 1.01 to 1.80, p=0.044, Table 1). The association between the GG genotype and the risk of AMI tended to be stronger among men who had no prior history of CHD (odds ratio 1.44, 95% CI 1.04 to 2.00, p=0.030) than among those who had prior CHD (odds ratio 1.32, 95% CI 0.81 to 2.17, p=0.314).

Other gene mutations that predicted AMI in the logistic model were the deletion/insertion in the alpha-2B-adrenergic receptor gene and the Thr98Ile SNP in the apolipoprotein B

gene (Table 1). Phenotypic data that added to the prediction of AMI were age, history of any atherosclerotic disease, cigarette-years of smoking, family history of CHD and diabetes, the presence of type 2 diabetes, and serum total and high-density lipoprotein (HDL) cholesterol (Table 1). Of these 12 variables, an empirical binary logistic function was constructed (Table 1). The population attributable risk, calculated across quintiles of the risk score, according to Miettinen OS, was 0.76. Odds ratios for quintiles (the lowest as reference): 12.8, 95% confidence interval (CI) 7.2 to 22.9, 6.4 (3.5 to 11.5), 2.4 (1.3 to 4.6) and 1.5 (0.8 to 3.1). When a split at the predicted probability (score value) of 0.2 was used, the odds ratio was 5.3 with 95% CI 4.0 to 7.0, $p<0.001$.

We also analyzed the prediction by gene mutations and phenotypic data the risk of AMI within five years of the baseline examination (Table 2). Another beta-defensin (DEFB129) SNP, located in IVS1-13_12insCTC, was a strong predictor of AMI. The carriers of the insertion CTC allele had 2.3-fold risk of AMI (95% CI 1.4 to 3.9, $p=0.002$). Also the apolipoprotein B Thr homozygosity predicted AMI strongly, and the deletion homogenicity of alpha-2B-adrenergic receptor gene fairly strongly. Phenotypic data that predicted AMI in five years were age, history of any atherosclerotic disease, cigarette-years of smoking, the presence of hypertension, the use of cholesterol lowering medication, family history of CHD and diabetes, waist-to-hip circumference ratio, and serum concentrations of total and high-density lipoprotein (HDL) cholesterol and ferritin. When the default split of the predicted probability (0.50) was used, the model predicted correctly 95.5% of the observed AMIs. When a split at the predicted probability of 0.2 was used, the odds ratio was 11.1 with 95% CI 5.9 to 21.2, $p<0.001$.

We also analyzed the predictors of AMI in men who had a family history of CHD (Table 3). The same three mutations predicted AMI. Of the measurements by questionaire the strongest predictors were the history of CHD in the subject and his socioeconomic status. Of the biochemical measurements, the most predictive were serum ferritin copncentration (classified into two categories), serum C-reactive protein, serum LDL cholesterol and serum HDL cholesterol (protective). When the default split of the predicted probability (0.50) was used, the model predicted correctly 94.0% of the observed AMIs. When a split at the predicted probability of 0.2 was used, the odds ratio was 8.2 with 95% CI 4.0 to 16.8, $p<0.001$.

In another statistical analysis, we analyzed the predictors of AMI within two years of risk factor measurements (Table 4). The Leu54Met mutation in the paraoxonase 1 gene and Cys282Tyr mutation in the HFE (HLA-H) gene were the strongest genetic predictors of AMI. Other, non-genetic predictors are presented in Table 4.

Thus, we disclose here a novel genetic test based on genotyping mutations in a human defensin gene, such as human beta-defensin 1 and 129 genes, with an optional multivariable model that predicts future myocardial infarction very well in the data set they were derived of. On the basis of our invention and empirical evidence supporting it, mutations in the human beta-defensins are associated with an increased risk of AMI and CHD both in healthy persons and in those who have a family history of CHD. Thus, for the first time it is showed that defensins are related to AMI and CHD and a mutation in a defensin gene can be a statistically significant risk factor for AMI and CHD.

When information of a few important mutations is combined with phenotypic information, the prediction of a multivariate risk prediction model is enhanced. An advantage is that only a small number of genotypings and biochemical or other measurements need to be carried out and a very short self-administered questionnaire needs to be filled in. The risk model can be estimated/constructed for different lenghts of follow-up, enabling the use of them for different purposes.

Table 1: A multivariate logistic model predicting the risk of MI in 1548 men in 9-15 years (256 experienced an AMI during the follow-up).

Predictor	Mutation	Coefficient (b_i)	S.E.	p-value	Odds ratio	95% Confidence interval
Beta-Defensin 1 (GG hozygote vs. other)	3'UTR+5 A/G	0.30	0.15	0.044	1.35	1.01, 1.80
Alpha-2B-adrenergic receptor (deletion carrier vs. non-carrier)	Insertion/deletion	0.56	0.21	0.007	1.75	1.16, 2.65
Alpha-2B-adrenergic receptor (I/D heterozygote vs. non-carrier)	Insertion/deletion	0.31	0.18	0.088	1.36	0.96, 1.94
Apolipoprotein B (Thr homozygote vs. other)	Thr98Ile (Thr71Ile)	0.49	0.27	0.067	1.63	0.97, 2.74
Age (per year)	NA	0.08	0.016	<0.001	1.08	1.05, 1.12
History of atherosclerotic disease (yes vs. no)	NA	0.69	0.16	<0.001	1.99	1.45, 2.72
Cigarette-years (per cigarettes/d multiplied by years smoked)	NA	0.001	<0.001	0.001	1.001	1.00, 1.001
CHD in the family (yes vs. no)	NA	0.64	0.15	<0.001	1.90	1.41, 2.56
Diabetes in the family (yes vs. no)	NA	0.48	0.16	0.002	1.62	1.19, 2.21
Diabetes in the subject (yes vs. no)	NA	1.23	0.29	<0.001	3.41	1.93, 6.04
Serum total cholesterol (per 1.0 mmol/L)	NA	0.22	0.07	0.001	1.25	1.09, 1.44
Serum HDL cholesterol (per 1.0 mmol/L)	NA	-0.84	0.27	0.002	0.43	0.26, 0.73

Constant 9.784.

Table 2: A multivariate logistic model predicting the risk of MI in 1548 men in 5 years (of whom 69 experienced an AMI during the follow-up).

Predictor	Mutation	Coefficient (b_i)	S.E.	p-value	Odds ratio	95% Confidence interval
Beta-Defensin 129 (insertion CTC carrier vs. other)	IVS1-13_12insCTC	0.831	0.268	0.002	2.30	1.36, 3.88
Alpha-2B-adrenergic receptor (deletion homozygote vs. other)	Insertion/deletion	0.295	0.298	0.321	1.34	0.75, 2.41
Apolipoprotein B (Thr homozygote vs. other)	Thr98Ile (Thr71Ile)	1.227	0.374	0.001	3.41	1.64, 7.10
Age (per year)	NA	0.078	0.032	0.016	1.081	1.02, 1.15
History of atherosclerotic disease (yes vs. no)	NA	0.766	0.277	0.006	2.15	1.25, 3.70
Cigarette-years (per cigarettes/d multiplied by years smoked) divided by 100	NA	0.072	0.035	0.037	1.08	1.004, 1.15
Hypertension (yes vs. no)	NA	0.447	0.277	0.107	1.56	0.91, 2.69
Waist-to-hip circumference ratio (m/cm)	NA	0.024	0.018	0.186	1.02	0.99, 1.06
CHD in the family (yes vs. no)	NA	0.843	0.285	0.003	2.32	1.33, 4.06
Diabetes in the family (yes vs. no)	NA	0.352	0.276	0.202	1.42	0.83, 2.44
Cholesterol lowering medication (yes vs. no)	NA	1.713	0.844	0.042	5.55	1.06, 29.0
Serum total cholesterol (per 1.0 mmol/L)	NA	0.174	0.118	0.143	1.19	0.94, 1.50
Serum HDL cholesterol (per 1.0 mmol/L)	NA	-0.818	0.512	0.110	0.44	0.16, 1.20
Serum ferritin (per 100 micrograms/L)	NA	0.131	0.062	0.034	1.14	1.01, 1.29

Constant term 14.144. For adjustment purpose, the model also included a term for examination month. The models predicted correctly 95.5% of the observed acute myocardial infarctions ($p<0.001$).

Table 3: A multivariate logistic model predicting the 5-year risk of MI in 761 men with a family history of CHD (of whom 49 experienced an AMI during the follow-up).

Predictor	Mutation	Coefficient (b_i)	S.E.	p-value	Odds ratio	95% Confidence interval
Beta-Defensin 129 (insertion CTC carrier vs. other)	IVS1-13_12insCTC	0.555	0.327	0.090	1.74	0.92, 3.30
Alpha-2B-adrenergic receptor (deletion homozygote vs. other)	Insertion/deletion	0.916	0.339	0.007	2.50	1.29, 4.86
Apolipoprotein B (Thr homozygote vs. other)	Thr98Ile (Thr71Ile)	1.151	0.469	0.014	3.16	1.26, 7.93
History of CHD (yes vs. no)	NA	1.220	0.337	<0.001	3.39	1.75, 6.56
Socioeconomic status (score of 0 to 23)	NA	0.079	0.042	0.063	1.08	1.00, 1.18
Hypertension (yes vs. no)	NA	0.492	0.328	0.134	1.64	0.86, 3.11
Serum LDL cholesterol (per 1.0 mmol/L)	NA	0.190	0.156	0.224	1.21	0.89, 1.64
Serum HDL cholesterol (per 1.0 mmol/L)	NA	-0.896	0.644	0.164	0.41	0.12, 1.44
Serum C-reactive protein (mg/L)		0.063	0.038	0.097	1.07	0.99, 1.15
Serum ferritin (>200 micrograms/L vs. less)	NA	1.017	0.336	0.002	2.77	1.43, 5.34

Constant term 17.612. For adjustment purpose, the model also included a term for examination month. The models predicted correctly 94.0% of the observed acute myocardial infarctions (p<0.001).

Table 4: A multivariate logistic model predicting the 2-year risk of MI in 1587 men (of whom 31 experienced an AMI during the follow-up).

Predictor	Mutation	Coefficient (b_i)	S.E.	p-value	Odds ratio	95% Confidence interval
Paraoxonase 1	Leu54Met	1.400	0.438	0.001	4.06	1.72, 9.57
HFE	Cys282Tyr	0.499	0.590	0.398	1.65	0.52, 5.24
History of prior AMI (yes vs. no)	NA	1.298	0.581	0.025	3.66	1.17, 11.43
History of claudication (yes vs. no)	NA	1.311	0.574	0.022	3.71	1.20, 11.44
Antihypertensive medication (yes vs. no)	NA	0.755	0.453	0.095	2.13	0.88, 5.16
Family history of cerebrovascular stroke (yes vs. no)	NA	0.894	0.411	0.030	2.45	1.09, 5.47
Waist-to-hip circumference ratio (m/cm)	NA	0.040	0.023	0.081	1.04	1.00, 1.09
Serum cholesterol (per 1.0 mmol/L)	NA	0.310	0.162	0.056	1.36	0.99, 1.87
Serum ferritin (>200 micrograms/L vs. less)	NA	0.932	0.393	0.018	2.54	1.18, 5.48

Constant term 30.575. For adjustment purpose, the model also included a term for examination month. The models predicted correctly 98.1% of the observed acute myocardial infarctions ($p<0.001$).

REFERENCES:

Barnathan ES, Raghunath PN, Tomaszewski JE, Ganz T, Cines DB, Higazi A al-R. Immunohistochemical localization of defensin in human coronary vessels. Am J Pathol. 1997; 150: 1009-20.

Bensch KW, Raida M, Magert HJ, Schulz-Knappe P, Forssmann WG. hBD-1: a novel beta-defensin from human plasma. FEBS Lett. 1995; 368: 331-5.Callen DF, Baker E, Simmers RN, Seshadri R, Roninson IB. Localization of the human multiple drug resistance gene, MDR1, to 7q21.1. Hum. Genet. 1987; 77: 142-144.

Diverse Populations Collaborative Group. Prediction of mortality from coronary heart disease among diverse populations: is there a common predictive function? Heart 2002; 88: 222-8.

Dork T, Stuhrmann M. Polymorphisms of the human beta-defensin-1 gene. Mol Cell Probes. 1998; 12: 171-3.

Ganz T, Lehrer RI. Defensins. Pharmacol Ther. 1995; 66: 191-205.

Hoover DM, Chertov O, Lubkowski J. The structure of human beta-defensin-1: new insights into structural properties of beta-defensins. J Biol Chem. 2001; 276: 39021-6.

Jia HP, Schutte BC, Schudy A, Linzmeier R, Guthmiller JM, Johnson GK, et al. Discovery of new human beta-defensins using a genomics-based approach. Gene. 2001; 263: 211-8.

Lehmann J, Retz M, Harder J, Krams M, Kellner U, Hartmann J, et al. Expression of human beta-defensins 1 and 2 in kidneys with chronic bacterial infection. BMC Infect Dis. 2002; 2: 20.

Nedelcheva Kristensen V, Kelefiotis D, Kristensen T and Borresen-Dale: High-Throughput methods for detection of genetic variation. Biotechniques 30:318-332, 2001.

Salonen JT. Is there a continuing need for longitudinal epidemiologic research - The Kuopio Ischaemic Heart Disease Risk Factor Study. Ann Clin Res 1988: 20: 46-50.

Salonen JT, Malin R, Tuomainen T-P, Nyyssönen K, Nissinen T, Lakka TA, Lehtimäki T. Polymorphism in the high density lipoprotein paraoxonase gene and the risk of acute myocardial infarction in men: a prospective population-based study. Brit Med J 1999; 319: 487-9.

Schutte BC, McCray PB Jr. β-defensins in lung host defense. Annu Rev Physiol. 2002; 64: 709-48.

Snapir A, Heinonen P, Tuomainen T-P, Alhopuro P, Karvonen MK, Lakka TA, Nyyssönen K, Salonen R, Kauhanen J, Valkonen V-P, Pesonen U, Koulu M, Scheinin M, Salonen JT. An Insertion/Deletion polymorphism in the α_{2B}-adrenergic receptor gene is a novel genetic risk factor for acute coronary events. J Am Coll Cardiol 2001; 37: 1516-1522.

Syvänen A-C: Accessing genetic variation: genotyping single nucleotide polymorphisms. Nature reviews/ Genetics 2:930-942, 2001

Tuomainen T-P, Kontula K, Nyyssönen K, Lakka TA, Heliö T, Salonen JT. Increased risk of acute myocardial infarction in carriers of the hemochromatosis gene Cys282Tyr mutation: A prospective cohort study in men in Eastern Finland. Circulation 1999; 100: 1274-1279.

Valore EV, Park CH, Quayle AJ, Wiles KR, McCray PB Jr, Ganz T. Human beta-defensin-1: an antimicrobial peptide of urogenital tissues. J Clin Invest. 1998; 101: 1633-42.

L3

Claims:

1. A method for detecting genetic variation or polymorphism, i.e. a mutation, in a defensin gene comprising the steps of:

   i) providing a biological sample taken from a subject to be tested,

   ii) detecting the presence or absence of a variant genotype of the defensin gene in the biological sample, the presence of a variant defensin genotype indicating an increased risk of cardiovascular disease in said subject.

2. The method according to claim 1, wherein said variant genotype of the defensin gene is a homo- or heterozygote form of the mutation.

3. The method according to claim 1, wherein the detection step is a DNA-assay.

4. The method according to claim 1, wherein the detection step is carried out using a gene or DNA chip, microarray, strip, panel or similar combination of more than one genes, mutations or RNA expressions to be assayed.

5. The method according to claim 1, wherein the allelic pattern is determined using polymerase chain reaction.

6. The method according to claim 1, wherein the biological sample is a blood sample or buccal swab sample and genomic DNA is isolated from said sample.

7. The method according to claim 1, wherein the detection step is based on a capturing probe.

8. The method according to claim 1, wherein said method is used for determining whether a subject will benefit from treatment with a drug, nutrient or other therapy enhancing the defensin production, levels or activity or inhibiting defensin catabolism or elimination in the subject.

9. The method according to claim 1, wherein said method is used for determining whether a subject will be at increased risk of adverse effects or reactions if defensin antagonists are administered to a subject.

10. The method according to claim 1, further comprising a step of selecting a subject with a defensin gene sequence reducing the expression, production or levels of defensin protein for clinical drug trials testing the anticoronary and myocardial ischaemia preventing effects of compounds.

11. The method according to claim 1, wherein said cardiovascular disease is acute myocardial infarction (AMI) or coronary heart disease (CHD).

12. The method according to any one of the previous claims, wherein said defensin is selected from the group consisting of: beta-defensin-1, beta-defensin-129, and alfa-defensin-5.

13. The method according to claim 12, wherein said variant genotype is human beta-defensin-1 gene comprising 3`UTR +5A→G mutation.

14. The method according to claim 12, wherein said variant genotype is human alfa-defensin-5 gene comprising IVS1 +198C→T mutation and/or IVS1 +243G→C mutation.

15. The method according to claim 12, wherein said variant genotype is human beta-defensin-129 gene comprising IVS1-13_12insCTC mutation.

16. The method according to claim 1, wherein genetic variation is further determined from the genes selected from the group consisting of:

a) alpha$_{2B}$-adrenoceptor,
b) apolipoprotein B, and
c) beta-2-adrenergic receptor

wherein the presence of a variant genotype in said genes indicates an increased risk of cardiovascular disease in said subject.

17. The method according to claim 16, wherein said variant genotype is alpha-$_{2B}$-adrenoceptor gene comprising insertion/deletion mutation, or said variant genotype is beta-2-adrenergic receptor comprising Gly16Arg and/or Glu27Gln mutation.

18. The method according to claim 16, wherein said variant genotype is apolipoprotein B gene comprising Thr98Ile mutation.

19. The method according to any one of the previous claims, further comprising a step of combining information concerning age, gender, the family history of hypertension, diabetes and hypercholesterolemia, and the medical history concerning cardiovascular diseases or diabetes of the subject with the results obtained from step ii) of the method for confirming the indication obtained from the detection step.

20. The method according to claim 19, wherein said information is about hypercholesterolemia in the family, smoking status, use of cholesterol lowering medications, CHD in the family, history of cardiovascular disease, obesity in the family, and waist-to-hip circumference ratio (cm/cm)

21. The method according to claim 19, further comprising a step determining blood, serum or plasma cholesterol, HDL cholesterol, LDL cholesterol, triglyceride, apolipoprotein B and AI, fibrinogen, ferritin, transferrin receptor, C-reactive protein, serum or plasma insulin concentration.

22. The method according to claim 19, wherein the detected mutations are 3'UTR+5 A/G of the beta-defensin-1 gene, an insertion/deletion of three glutamic acids in the region of 12 Glu aminoacids in the codons 298-309 of Alpha-$_{2B}$-adrenoceptor, and the Thr98Ile of apolipoprotein B gene.

23. The method according to claim 19 further comprising a step of calculating the probability of a cardiovascular disease using a logistic regression equation as follows:

Probability of a cardiovascular disease = $[1 + e^{(-(-a + \Sigma(bi*Xi))}]^{-1}$, where e is Napier's constant, X_i are variables related to the cardiovascular disease, b_i are coefficients of these variables in the logistic function, and a is the constant term in the logistic function.

24. The method according to claim 23, wherein a and b_i are determined in the population in which the method is to be used.

25. The method according to claim 23, wherein Xi are selected among the variables that have been measured in the population in which the method is to be used.

26. The method according to claim 23, wherein b_i are between the values of –20 and 20 and/or wherein X_i are binary variables that can have values or are coded as 0 (zero) or 1 (one).

27. The method according to claim 23, wherein i are between the values 0 (none) and 100,000.

28. The method according to claim 23, wherein subject's short term, median term, and/or long term risk of CHD and/or AMI is predicted.

29. A method for targeting the treatment of CHD and AMI in a subject with CHD by determining the pattern of alleles encoding a variant defensin, i.e. by determining if said subject's genotype of the defensin is of the variant type, comprising the steps presented in claim 1, and treating a subject of the variant genotype with a drug affecting defensin production or metabolism of the subject.

30. The method according to claim 29, wherein said defensin is as defined in claim 12.

31. The method according to claim 30, wherein the variant genotype is as defined in any one of claims 13-15.

32. The method according to claim any one of calims 29-31, wherein said variant genotype of the defensin is a homozygote or heterozygote form of mutation.

33. The method according to claims 29, wherein said CHD is angina pectoris or other form of CHD.

34. A method for treating a human or animal suffering from CHD or AMI, said method comprising a therapy enhancing defensin availability, production or concentration of the human subject or animal.

35. The method of claim 34, wherein said animal is a mammal.

36. A method for treating vascular complications of CHD and AMI, said method comprising a step of enhancing defensin availability, production or concentration in the circulation of a human subject or animal.

37. The method according to claim 34 or 36, wherein said defensin is as defined in claim 12.

38. The method according to claim 37, said method comprising administering to a subject a compound enhancing Beta-defensin-1 availability, production or concentration of the subject.

39. The method according to claim 34 or 36, wherein the said method of treating is a dietary treatment or a vaccination.

40. The method according to claim 34 or 20, wherein said therapy is gene therapy or gene transfer.

41. The method according to claim 40, wherein said therapy comprises the transfer of the non-variant Beta-defensin 1 gene or fragment or derivative thereof.

42. A kit for detecting genetic variation or polymorphism, i.e. a mutation, in a defensin gene for the determination of a risk of acute myocardial infarction, AMI, and coronary heart disease, CHD, in a subject, comprising means for defensin gene allele detection, and optionally software to interpret the results of the determination.

43. The kit according to claim 42, wherein said defensin is as defined in claim 12.

44. The kit according to claim 42, wherein genetic variation or polymorphism, i.e. a mutation, is further detected in the genes selected from the group consisting of:

a) alpha$_{-2B}$-adrenoceptor
b) apolipoprotein B, and
c) beta-2-adrenergic receptor.

45. The method according to claim 44, wherein the genetic variation to be detected is as defined in any one of claims 13-15.

46. The kit according to claim 45 comprising a capturing nucleic acid probe specifically binding to the variant genotype as defined in any one of claims 13-15 .

47. The kit according to any one of claims 42-46, comprising a DNA chip, microarray, DNA strip, DNA panel or real-time PCR based tests.

48. The kit according to any one of claims 42-47 , comprising a questionnaire for obtaining patient information concerning age, gender, height, weight, the family history of hypertension and hypercholesterolemia, the medical history concerning cardiovascular diseases.

49. An isolated variant nucleic acid encoding alfa-defensin-5 protein, said nucleic acid comprising IVS1 +198C→T and/or IVS1 +243G→C mutation.

50. An isolated variant nucleic acid encoding beta-defensin-129 protein, said nucleic acid comprising IVS1 –13_12 in/del CTC mutation.

51. The nucleic acid according to claim 49 or 50, wherein said nucleic acid is a genomic nucleotide sequence.

52. The nucleic acid according to claim 51, wherein said nucleic acid is cDNA.

53. The nucleic acid according to claim 49 or 50 comprising an RNA sequence.

54. The nucleic acid according to 49 having the nucleic acid sequence set forth in SEQ ID NO:7.

55. The nucleic acid according to 50 having the nucleic acid sequence set forth in SEQ ID NO:32.

56. A capturing probe binding to the nucleic acid according to claim 49 or 50.

57. The capturing probe according to claim 56, which comprises a single strand of the cDNA according to claim 52.

58. The capturing probe according to claim 56 or 57, which is specifically binding to variant defensin nucleic acid according to claim 49 or 50, but do not bind non-variant defensin.

59. A method for determining the presence or absence of a nucleic acid as defined in claim 49 or 50 in a biological sample comprising the steps of:

a) treating said sample to obtain single stranded target nucleic acid, or if the target nucleic acid are already single stranded, directly employing step (b);
b) contacting said target nucleic acid with a capturing nucleic acid probe and a detector nucleic acid probe;

c) detecting the complex of capturing probe, target nucleic acid and detector probe.

60. The method according to claim 59, wherein the capturing nucleic acid probe is attached or capable of attaching to a solid phase, and comprises the cDNA sequence according to claim 52, and wherein a detected signal from the solid phase is an indication of the presence in the sample of a nucleic acid as defined in claim 49 or 50.

61. The method according to claim 60, wherein the capturing nucleic acid probe is attached or capable of attaching to a solid phase, and comprises a cDNA corresponding to the gene coding a wild-type defensin protein, and wherein a detected signal from the solid phase is an indication of the absence of the nucleic acid as defined in claim 49 or 50 in the sample.

62. A transgenic animal which carries a human DNA sequence comprising a nucleotide sequence encoding a variant defensin nucleic acid as defined in claim 49 or 50.

63. RNA interference methods and models involving a variant nucleotide sequence encoding a variant defensin nucleic acid as defined in claim 49 or 50.

64. A method for measuring defensin protein expression, production or concentration in human tissues, comprising the steps of:

a) providing a tissue sample taken from a subject to be tested,

b) detecting the expression, production or concentration of defensin protein in said sample, wherein reduced expression, production or concentration indicates an increased risk of cardiovascular disease in said subject.

L4

(57) Abstract

The present invention relates to a variant defensin gene. The invention provides a method of identifying subject's susceptibility or predisposition to or risk of developing myocardial infarction (MI) or coronary heart disease (CHD) by detecting gene polymorphisms and other gene mutations from a biological sample of the subject and obtaining information concerning the family and medical history, blood, serum, plasma and urinary analytes of the subject. The invention also provides a multivariate model, a combination or algorithm of variables which best describes the probability of CHD, especially MI. The invention also relates to a test kit and software for accomplishing the method.

L5

SEQUENCE LISTING

<110> Oy Jurilab Ltd

<120> Method for detecting the risk of acute myocardial infarction and coronary heart disease

<130> 40597

<160> 56

<170> PatentIn version 3.1

<210> 1
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR primer
<400> 1
cataatttca gcccgatgtg 20

<210> 2
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR primer
<400> 2
caccctaacc ccctacttct 20

<210> 3
<211> 18
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR primer
<400> 3
gggcttgctc tttctttc 18

<210> 4
<211> 18
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR primer

<400> 4
tccttggttc ctctcatc 18

<210> 5
<211> 18
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR primer
<400> 5
ctgagtgtgc aggacgag 18

<210> 6
<211> 18
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR primer
<400> 6
cacattgcca aacacgat 18

<210> 7
<211> 736
<212> DNA
<213> Homo sapiens
<400> 7
agaaagagga gcatcaaagg gatcttgaga acaaaggcag tccttcccct cccaatcaca 60

tgcccacctc ctctcactgc agcttctgtc tcaggtcttc tcccagcaga gctataaatc 120

caggctgact cctcactccc cacatatcca ctcctgctct ccctcctgca ggtgacccca 180

gccatgagga ccatcgccat ccttgctgcc attctcctgg tggccctgca ggcccaggct 240

gagtcactcc aggaaagagc tgatgaggct acaacccaga agcagtctgg ggaagacaac 300

caggaccttg ctatctcctt tgcaggaaat ggactctctg ctcttagaac ctcaggtagg 360

agacatcaat cttgcacatc tgcaaaatct agaaaaaaag gattggagaa aggatctgga 420

gtcaagtgtg gaaaggtcta cctcacttga gtgactttac ttaatcttcc tggaccttga 480

ttttctcatc tataaattaa tcagtgagaa ccaaataaat ctaaaagatt ttcttttttc 540

taagactttc agttccaaga tatttctgtg aaatttgcta cttttaagat agaaagacct 600

acactgacta gttctttgta gatctaaatg ggcagactta gttatataga gagtgtttta 660

ctttgtccat tggaaaagct tttagaacct agagaggaac ctataggtgt gttttgatgt 720

aggctaatag gcttga 736

<210> 8
<211> 736
<212> DNA
<213> Homo sapiens
<400> 8

agaaagagga gcatcaaagg gatcttgaga acaaaggcag tccttcccct cccaatcaca 60

tgcccacctc ctctcactgc agcttctgtc tcaggtcttc tcccagcaga gctataaatc 120

caggctgact cctcactccc cacatatcca ctcctgctct ccctcctgca ggtgacccca 180

gccatgagga ccatcgccat ccttgctgcc attctcctgg tggccctgca ggcccaggct 240

gagtcactcc aggaaagagc tgatgaggct acaacccaga agcagtctgg ggaagacaac 300

caggaccttg ctatctcctt tgcaggaaat ggactctctg ctcttagaac ctcaggtagg 360

agacatcaat cttgcacatc tgcaaaatct agaaaaaaag gattggagaa aggatctgga 420

gtcaagtgtg gaaaggtcta cctcacttga gtgactttac ttaatcttcc tggaccttga 480

ttttctcatc tataaattaa tcagtgagaa ccaaataaat ctaaaagatt ttcttttttc 540

taagactttc agctccaaga tatttctgtg aaatttgcta cttttaagat agaaagagct 600

acactgacta gttctttgta gatctaaatg ggcagactta gttatataga gagtgtttta 660

ctttgtccat tggaaaagct tttagaacct agagaggaac ctataggtgt gttttgatgt 720

aggctaatag gcttga 736

<210> 9
<211> 19
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR primer
<400> 9
agaaagagga gcatcaaag 19

<210> 10
<211> 19
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR primer
<400> 10
tcaagcctat tagcctaca 19

<210> 11
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR primer
<400> 11
gacaacctca atgctctgct 20

<210> 12
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer
<400> 12
tgacttacct ggacatggct 20

<210> 13
<211> 35
<212> DNA
<213> Artificial Sequence
<220>
<223> Snapshot primer
<400> 13
ttttttttt tttttctttt ttctaagact ttcag 35

<210> 14
<211> 40
<212> DNA
<213> Artificial Sequence
<220>
<223> Snapshot primer
<400> 14
tttttttttt tttttttttg ctacttttaa gatagaaaga 40

<210> 15
<211> 45
<212> DNA
<213> Artificial Sequence
<220>
<223> Snapshot primer
<400> 15
tttttttttt tttttttttt tttttttagt gctgcaagtg agctg 45

<210> 16
<211> 50
<212> DNA
<213> Artificial Sequence
<220>
<223> Snapshot primer
<400> 16
tttttttttt tttttttttt tttttttttt tttccagaga ggaagccttg 50

<210> 17
<211> 55
<212> DNA
<213> Artificial Sequence
<220>
<223> Snapshot primer
<400> 17
tttttttttt tttttttttt tttttttttt tttttttttc ttgctggcac ccaat 55

<210> 18
<211> 60
<212> DNA
<213> Artificial Sequence
<220>

<223> Snapshot primer
<400> 18
tttttttttt tttttttttt tttttttttt tttttttttt tttaccacga cgtcacgcag 60

<210> 19
<211> 30
<212> DNA
<213> Artificial Sequence
<220>
<223> Snapshot primer
<400> 19
tttttttttt tttgaagacc agccagtgca 30

<210> 20
<211> 1344
<212> DNA
<213> Homo sapiens
<220>
<221> CDS
<222> (1)..(1344)
<223> Coding sequence for variant human ADRA2B gene
<400> 20
atg gac cac cag gac ccc tac tcc gtg cag gcc aca gcg gcc ata gcg 48
Met Asp His Gln Asp Pro Tyr Ser Val Gln Ala Thr Ala Ala Ile Ala
1 5 10 15

gcg gcc atc acc ttc ctc att ctc ttt acc atc ttc ggc aac gct ctg 96
Ala Ala Ile Thr Phe Leu Ile Leu Phe Thr Ile Phe Gly Asn Ala Leu
20 25 30

gtc atc ctg gct gtg ttg acc agc cgc tcg ctg cgc gcc cct cag aac 144
Val Ile Leu Ala Val Leu Thr Ser Arg Ser Leu Arg Ala Pro Gln Asn
35 40 45

ctg ttc ctg gtg tcg ctg gcc gcc gcc gac atc ctg gtg gcc acg ctc 192
Leu Phe Leu Val Ser Leu Ala Ala Ala Asp Ile Leu Val Ala Thr Leu
50 55 60

atc atc cct ttc tcg ctg gcc aac gag ctg ctg ggc tac tgg tac ttc 240
Ile Ile Pro Phe Ser Leu Ala Asn Glu Leu Leu Gly Tyr Trp Tyr Phe
65 70 75 80

cgg cgc acg tgg tgc gag gtg tac ctg gcg ctc gac gtg ctc ttc tgc 288
Arg Arg Thr Trp Cys Glu Val Tyr Leu Ala Leu Asp Val Leu Phe Cys
85 90 95

acc tcg tcc atc gtg cac ctg tgc gcc atc agc ctg gac cgc tac tgg 336
Thr Ser Ser Ile Val His Leu Cys Ala Ile Ser Leu Asp Arg Tyr Trp
100 105 110

gcc gtg agc cgc gcg ctg gag tac aac tcc aag cgc acc ccg cgc cgc 384
Ala Val Ser Arg Ala Leu Glu Tyr Asn Ser Lys Arg Thr Pro Arg Arg
115 120 125

atc aag tgc atc atc ctc act gtg tgg ctc atc gcc gcc gtc atc tcg 432
Ile Lys Cys Ile Ile Leu Thr Val Trp Leu Ile Ala Ala Val Ile Ser
130 135 140

ctg ccg ccc ctc atc tac aag ggc gac cag ggc ccc cag ccg cgc ggg 480

Leu Pro Pro Leu Ile Tyr Lys Gly Asp Gln Gly Pro Gln Pro Arg Gly
145 150 155 160

cgc ccc cag tgc aag ctc aac cag gag gcc tgg tac atc ctg gcc tcc 528
Arg Pro Gln Cys Lys Leu Asn Gln Glu Ala Trp Tyr Ile Leu Ala Ser
165 170 175

agc atc gga tct ttc ttt gct cct tgc ctc atc atg atc ctt gtc tac 576
Ser Ile Gly Ser Phe Phe Ala Pro Cys Leu Ile Met Ile Leu Val Tyr
180 185 190

ctg cgc atc tac ctg atc gcc aaa cgc agc aac cgc aga ggt ccc agg 624
Leu Arg Ile Tyr Leu Ile Ala Lys Arg Ser Asn Arg Arg Gly Pro Arg
195 200 205

gcc aag ggg ggg cct ggg cag ggt gag tcc aag cag ccc cga ccc gac 672
Ala Lys Gly Gly Pro Gly Gln Gly Glu Ser Lys Gln Pro Arg Pro Asp
210 215 220

cat ggt ggg gct ttg gcc tca gcc aaa ctg cca gcc ctg gcc tct gtg 720
His Gly Gly Ala Leu Ala Ser Ala Lys Leu Pro Ala Leu Ala Ser Val
225 230 235 240

gct tct gcc aga gag gtc aac gga cac tcg aag tcc act ggg gag aag 768
Ala Ser Ala Arg Glu Val Asn Gly His Ser Lys Ser Thr Gly Glu Lys
245 250 255

gag gag ggg gag acc cct gaa gat act ggg acc cgg gcc ttg cca ccc 816
Glu Glu Gly Glu Thr Pro Glu Asp Thr Gly Thr Arg Ala Leu Pro Pro
260 265 270

agt tgg gct gcc ctt ccc aac tca ggc cag ggc cag aag gag ggt gtt 864
Ser Trp Ala Ala Leu Pro Asn Ser Gly Gln Gly Gln Lys Glu Gly Val
275 280 285

tgt ggg gca tct cca gag gat gaa gct gaa gag gag gaa gag gag gag 912
Cys Gly Ala Ser Pro Glu Asp Glu Ala Glu Glu Glu Glu Glu Glu Glu
290 295 300

gag gag tgt gaa ccc cag gca gtg cca gtg tct ccg gcc tca gct tgc 960
Glu Glu Cys Glu Pro Gln Ala Val Pro Val Ser Pro Ala Ser Ala Cys
305 310 315 320

agc ccc ccg ctg cag cag cca cag ggc tcc cgg gtg ctg gcc acc cta 1008
Ser Pro Pro Leu Gln Gln Pro Gln Gly Ser Arg Val Leu Ala Thr Leu
325 330 335

cgt ggc cag gtg ctc ctg ggc agg ggc gtg ggt gct ata ggt ggg cag 1056
Arg Gly Gln Val Leu Leu Gly Arg Gly Val Gly Ala Ile Gly Gly Gln
340 345 350

tgg tgg cgt cga agg gcg cac gtg acc cgg gag aag cgc ttc acc ttc 1104
Trp Trp Arg Arg Arg Ala His Val Thr Arg Glu Lys Arg Phe Thr Phe
355 360 365

gtg ctg gct gtg gtc att ggc gtt ttt gtg ctc tgc tgg ttc ccc ttc 1152
Val Leu Ala Val Val Ile Gly Val Phe Val Leu Cys Trp Phe Pro Phe
370 375 380

ttc ttc agc tac agc ctg ggc gcc atc tgc ccg aag cac tgc aag gtg 1200
Phe Phe Ser Tyr Ser Leu Gly Ala Ile Cys Pro Lys His Cys Lys Val

```
385                 390                 395                 400

ccc cat ggc ctc ttc cag ttc ttc ttc tgg atc ggc tac tgc aac agc      1248
Pro His Gly Leu Phe Gln Phe Phe Phe Trp Ile Gly Tyr Cys Asn Ser
                405                 410                 415

tca ctg aac cct gtt atc tac acc atc ttc aac cag gac ttc cgc cgt      1296
Ser Leu Asn Pro Val Ile Tyr Thr Ile Phe Asn Gln Asp Phe Arg Arg
            420                 425                 430

gcc ttc cgg agg atc ctg tgc cgc ccg tgg acc cag acg gcc tgg tga      1344
Ala Phe Arg Arg Ile Leu Cys Arg Pro Trp Thr Gln Thr Ala Trp
        435                 440                 445


&lt;210&gt;  21
&lt;211&gt;  447
&lt;212&gt;  PRT
&lt;213&gt;  Homo sapiens
&lt;400&gt;  21

Met Asp His Gln Asp Pro Tyr Ser Val Gln Ala Thr Ala Ala Ile Ala
1               5                   10                  15

Ala Ala Ile Thr Phe Leu Ile Leu Phe Thr Ile Phe Gly Asn Ala Leu
            20                  25                  30

Val Ile Leu Ala Val Leu Thr Ser Arg Ser Leu Arg Ala Pro Gln Asn
        35                  40                  45

Leu Phe Leu Val Ser Leu Ala Ala Ala Asp Ile Leu Val Ala Thr Leu
    50                  55                  60

Ile Ile Pro Phe Ser Leu Ala Asn Glu Leu Leu Gly Tyr Trp Tyr Phe
65                  70                  75                  80

Arg Arg Thr Trp Cys Glu Val Tyr Leu Ala Leu Asp Val Leu Phe Cys
                85                  90                  95

Thr Ser Ser Ile Val His Leu Cys Ala Ile Ser Leu Asp Arg Tyr Trp
            100                 105                 110

Ala Val Ser Arg Ala Leu Glu Tyr Asn Ser Lys Arg Thr Pro Arg Arg
        115                 120                 125

Ile Lys Cys Ile Ile Leu Thr Val Trp Leu Ile Ala Ala Val Ile Ser
    130                 135                 140

Leu Pro Pro Leu Ile Tyr Lys Gly Asp Gln Gly Pro Gln Pro Arg Gly
145                 150                 155                 160
```

Arg Pro Gln Cys Lys Leu Asn Gln Glu Ala Trp Tyr Ile Leu Ala Ser
165 170 175

Ser Ile Gly Ser Phe Phe Ala Pro Cys Leu Ile Met Ile Leu Val Tyr
180 185 190

Leu Arg Ile Tyr Leu Ile Ala Lys Arg Ser Asn Arg Arg Gly Pro Arg
195 200 205

Ala Lys Gly Gly Pro Gly Gln Gly Glu Ser Lys Gln Pro Arg Pro Asp
210 215 220

His Gly Gly Ala Leu Ala Ser Ala Lys Leu Pro Ala Leu Ala Ser Val
225 230 235 240

Ala Ser Ala Arg Glu Val Asn Gly His Ser Lys Ser Thr Gly Glu Lys
245 250 255

Glu Glu Gly Glu Thr Pro Glu Asp Thr Gly Thr Arg Ala Leu Pro Pro
260 265 270

Ser Trp Ala Ala Leu Pro Asn Ser Gly Gln Gly Gln Lys Glu Gly Val
275 280 285

Cys Gly Ala Ser Pro Glu Asp Glu Ala Glu Glu Glu Glu Glu Glu Glu
290 295 300

Glu Glu Cys Glu Pro Gln Ala Val Pro Val Ser Pro Ala Ser Ala Cys
305 310 315 320

Ser Pro Pro Leu Gln Gln Pro Gln Gly Ser Arg Val Leu Ala Thr Leu
325 330 335

Arg Gly Gln Val Leu Leu Gly Arg Gly Val Gly Ala Ile Gly Gly Gln
340 345 350

Trp Trp Arg Arg Arg Ala His Val Thr Arg Glu Lys Arg Phe Thr Phe
355 360 365

Val Leu Ala Val Val Ile Gly Val Phe Val Leu Cys Trp Phe Pro Phe
370 375 380

Phe Phe Ser Tyr Ser Leu Gly Ala Ile Cys Pro Lys His Cys Lys Val
385 390 395 400

Pro His Gly Leu Phe Gln Phe Phe Phe Trp Ile Gly Tyr Cys Asn Ser

405 410 415

Ser Leu Asn Pro Val Ile Tyr Thr Ile Phe Asn Gln Asp Phe Arg Arg
420 425 430

Ala Phe Arg Arg Ile Leu Cys Arg Pro Trp Thr Gln Thr Ala Trp
435 440 445

<210> 22
<211> 1353
<212> DNA
<213> Homo sapiens
<220>
<221> CDS
<222> (1)..(1353)
<223> Coding sequence for human ADRA2B gene
<400> 22

atg gac cac cag gac ccc tac tcc gtg cag gcc aca gcg gcc ata gcg 48
Met Asp His Gln Asp Pro Tyr Ser Val Gln Ala Thr Ala Ala Ile Ala
1 5 10 15

gcg gcc atc acc ttc ctc att ctc ttt acc atc ttc ggc aac gct ctg 96
Ala Ala Ile Thr Phe Leu Ile Leu Phe Thr Ile Phe Gly Asn Ala Leu
20 25 30

gtc atc ctg gct gtg ttg acc agc cgc tcg ctg cgc gcc cct cag aac 144
Val Ile Leu Ala Val Leu Thr Ser Arg Ser Leu Arg Ala Pro Gln Asn
35 40 45

ctg ttc ctg gtg tcg ctg gcc gcc gcc gac atc ctg gtg gcc acg ctc 192
Leu Phe Leu Val Ser Leu Ala Ala Ala Asp Ile Leu Val Ala Thr Leu
50 55 60

atc atc cct ttc tcg ctg gcc aac gag ctg ctg ggc tac tgg tac ttc 240
Ile Ile Pro Phe Ser Leu Ala Asn Glu Leu Leu Gly Tyr Trp Tyr Phe
65 70 75 80

cgg cgc acg tgg tgc gag gtg tac ctg gcg ctc gac gtg ctc ttc tgc 288
Arg Arg Thr Trp Cys Glu Val Tyr Leu Ala Leu Asp Val Leu Phe Cys
85 90 95

acc tcg tcc atc gtg cac ctg tgc gcc atc agc ctg gac cgc tac tgg 336
Thr Ser Ser Ile Val His Leu Cys Ala Ile Ser Leu Asp Arg Tyr Trp
100 105 110

gcc gtg agc cgc gcg ctg gag tac aac tcc aag cgc acc ccg cgc cgc 384
Ala Val Ser Arg Ala Leu Glu Tyr Asn Ser Lys Arg Thr Pro Arg Arg
115 120 125

atc aag tgc atc atc ctc act gtg tgg ctc atc gcc gcc gtc atc tcg 432
Ile Lys Cys Ile Ile Leu Thr Val Trp Leu Ile Ala Ala Val Ile Ser
130 135 140

ctg ccg ccc ctc atc tac aag ggc gac cag ggc ccc cag ccg cgc ggg 480
Leu Pro Pro Leu Ile Tyr Lys Gly Asp Gln Gly Pro Gln Pro Arg Gly
145 150 155 160

cgc ccc cag tgc aag ctc aac cag gag gcc tgg tac atc ctg gcc tcc 528

```
Arg Pro Gln Cys Lys Leu Asn Gln Glu Ala Trp Tyr Ile Leu Ala Ser
                165                 170                 175

agc atc gga tct ttc ttt gct cct tgc ctc atc atg atc ctt gtc tac      576
Ser Ile Gly Ser Phe Phe Ala Pro Cys Leu Ile Met Ile Leu Val Tyr
            180                 185                 190

ctg cgc atc tac ctg atc gcc aaa cgc agc aac cgc aga ggt ccc agg      624
Leu Arg Ile Tyr Leu Ile Ala Lys Arg Ser Asn Arg Arg Gly Pro Arg
        195                 200                 205

gcc aag ggg ggg cct ggg cag ggt gag tcc aag cag ccc cga ccc gac      672
Ala Lys Gly Gly Pro Gly Gln Gly Glu Ser Lys Gln Pro Arg Pro Asp
    210                 215                 220

cat ggt ggg gct ttg gcc tca gcc aaa ctg cca gcc ctg gcc tct gtg      720
His Gly Gly Ala Leu Ala Ser Ala Lys Leu Pro Ala Leu Ala Ser Val
225                 230                 235                 240

gct tct gcc aga gag gtc aac gga cac tcg aag tcc act ggg gag aag      768
Ala Ser Ala Arg Glu Val Asn Gly His Ser Lys Ser Thr Gly Glu Lys
                245                 250                 255

gag gag ggg gag acc cct gaa gat act ggg acc cgg gcc ttg cca ccc      816
Glu Glu Gly Glu Thr Pro Glu Asp Thr Gly Thr Arg Ala Leu Pro Pro
            260                 265                 270

agt tgg gct gcc ctt ccc aac tca ggc cag ggc cag aag gag ggt gtt      864
Ser Trp Ala Ala Leu Pro Asn Ser Gly Gln Gly Gln Lys Glu Gly Val
        275                 280                 285

tgt ggg gca tct cca gag gat gaa gct gaa gag gag gaa gag gag gag      912
Cys Gly Ala Ser Pro Glu Asp Glu Ala Glu Glu Glu Glu Glu Glu Glu
    290                 295                 300

gag gag gag gaa gag tgt gaa ccc cag gca gtg cca gtg tct ccg gcc      960
Glu Glu Glu Glu Glu Cys Glu Pro Gln Ala Val Pro Val Ser Pro Ala
305                 310                 315                 320

tca gct tgc agc ccc ccg ctg cag cag cca cag ggc tcc cgg gtg ctg     1008
Ser Ala Cys Ser Pro Pro Leu Gln Gln Pro Gln Gly Ser Arg Val Leu
                325                 330                 335

gcc acc cta cgt ggc cag gtg ctc ctg ggc agg ggc gtg ggt gct ata     1056
Ala Thr Leu Arg Gly Gln Val Leu Leu Gly Arg Gly Val Gly Ala Ile
            340                 345                 350

ggt ggg cag tgg tgg cgt cga agg gcg cac gtg acc cgg gag aag cgc     1104
Gly Gly Gln Trp Trp Arg Arg Arg Ala His Val Thr Arg Glu Lys Arg
        355                 360                 365

ttc acc ttc gtg ctg gct gtg gtc att ggc gtt ttt gtg ctc tgc tgg     1152
Phe Thr Phe Val Leu Ala Val Val Ile Gly Val Phe Val Leu Cys Trp
    370                 375                 380

ttc ccc ttc ttc ttc agc tac agc ctg ggc gcc atc tgc ccg aag cac     1200
Phe Pro Phe Phe Phe Ser Tyr Ser Leu Gly Ala Ile Cys Pro Lys His
385                 390                 395                 400

tgc aag gtg ccc cat ggc ctc ttc cag ttc ttc ttc tgg atc ggc tac     1248
Cys Lys Val Pro His Gly Leu Phe Gln Phe Phe Phe Trp Ile Gly Tyr
```

```
                405                 410                 415

tgc aac agc tca ctg aac cct gtt atc tac acc atc ttc aac cag gac      1296
Cys Asn Ser Ser Leu Asn Pro Val Ile Tyr Thr Ile Phe Asn Gln Asp
            420                 425                 430

ttc cgc cgt gcc ttc cgg agg atc ctg tgc cgc ccg tgg acc cag acg      1344
Phe Arg Arg Ala Phe Arg Arg Ile Leu Cys Arg Pro Trp Thr Gln Thr
        435                 440                 445

gcc tgg tga                                                          1353
Ala Trp
    450


<210>  23
<211>  450
<212>  PRT
<213>  Homo sapiens
<400>  23

Met Asp His Gln Asp Pro Tyr Ser Val Gln Ala Thr Ala Ala Ile Ala
1               5                   10                  15

Ala Ala Ile Thr Phe Leu Ile Leu Phe Thr Ile Phe Gly Asn Ala Leu
            20                  25                  30

Val Ile Leu Ala Val Leu Thr Ser Arg Ser Leu Arg Ala Pro Gln Asn
        35                  40                  45

Leu Phe Leu Val Ser Leu Ala Ala Ala Asp Ile Leu Val Ala Thr Leu
    50                  55                  60

Ile Ile Pro Phe Ser Leu Ala Asn Glu Leu Leu Gly Tyr Trp Tyr Phe
65                  70                  75                  80

Arg Arg Thr Trp Cys Glu Val Tyr Leu Ala Leu Asp Val Leu Phe Cys
                85                  90                  95

Thr Ser Ser Ile Val His Leu Cys Ala Ile Ser Leu Asp Arg Tyr Trp
            100                 105                 110

Ala Val Ser Arg Ala Leu Glu Tyr Asn Ser Lys Arg Thr Pro Arg Arg
        115                 120                 125

Ile Lys Cys Ile Ile Leu Thr Val Trp Leu Ile Ala Ala Val Ile Ser
    130                 135                 140

Leu Pro Pro Leu Ile Tyr Lys Gly Asp Gln Gly Pro Gln Pro Arg Gly
145                 150                 155                 160
```

Arg Pro Gln Cys Lys Leu Asn Gln Glu Ala Trp Tyr Ile Leu Ala Ser
165 170 175

Ser Ile Gly Ser Phe Phe Ala Pro Cys Leu Ile Met Ile Leu Val Tyr
180 185 190

Leu Arg Ile Tyr Leu Ile Ala Lys Arg Ser Asn Arg Arg Gly Pro Arg
195 200 205

Ala Lys Gly Gly Pro Gly Gln Gly Glu Ser Lys Gln Pro Arg Pro Asp
210 215 220

His Gly Gly Ala Leu Ala Ser Ala Lys Leu Pro Ala Leu Ala Ser Val
225 230 235 240

Ala Ser Ala Arg Glu Val Asn Gly His Ser Lys Ser Thr Gly Glu Lys
245 250 255

Glu Glu Gly Glu Thr Pro Glu Asp Thr Gly Thr Arg Ala Leu Pro Pro
260 265 270

Ser Trp Ala Ala Leu Pro Asn Ser Gly Gln Gly Gln Lys Glu Gly Val
275 280 285

Cys Gly Ala Ser Pro Glu Asp Glu Ala Glu Glu Glu Glu Glu Glu Glu
290 295 300

Glu Glu Glu Glu Glu Cys Glu Pro Gln Ala Val Pro Val Ser Pro Ala
305 310 315 320

Ser Ala Cys Ser Pro Pro Leu Gln Gln Pro Gln Gly Ser Arg Val Leu
325 330 335

Ala Thr Leu Arg Gly Gln Val Leu Leu Gly Arg Gly Val Gly Ala Ile
340 345 350

Gly Gly Gln Trp Trp Arg Arg Arg Ala His Val Thr Arg Glu Lys Arg
355 360 365

Phe Thr Phe Val Leu Ala Val Val Ile Gly Val Phe Val Leu Cys Trp
370 375 380

Phe Pro Phe Phe Phe Ser Tyr Ser Leu Gly Ala Ile Cys Pro Lys His
385 390 395 400

Cys Lys Val Pro His Gly Leu Phe Gln Phe Phe Phe Trp Ile Gly Tyr

405 410 415

Cys Asn Ser Ser Leu Asn Pro Val Ile Tyr Thr Ile Phe Asn Gln Asp
420 425 430

Phe Arg Arg Ala Phe Arg Arg Ile Leu Cys Arg Pro Trp Thr Gln Thr
435 440 445

Ala Trp
450

<210> 24
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR primer
<400> 24
gggtgtttgt ggggcatctc 20

<210> 25
<211> 19
<212> DNA
<213> Artificial Sequence
<220>
<223> Snapshot primer
<400> 25
tggcactgcc tggggttca 19

<210> 26
<211> 18
<212> DNA
<213> Artificial Sequence
<220>
<223> Sequencing primer
<400> 26
tcaggtcttc tcccagca 18

<210> 27
<211> 619
<212> DNA
<213> Homo sapiens
<400> 27
ggatgaagca gaatgaagag taggtaaccc tgaggttgag aggtatattg ttggaccagg 60

gagcaggtaa taaatacatc ctggatagac tcacatgggg aaaaaaacta tgatcttgca 120

tgactaacac atagctagta agatttcttg tcacttacga caaagacatg aattttctcc 180

atcctaacat gactgataca gtgtctctta tttagactat ctcagttagt ctggctgtgc 240

ttgtcctttt tcccacctcc ctcgctgtgc ctgaccctct cttctttcca caggttctca 300

ggcaagagcc acctgctatt gccgaaccgg ccgttgtgct acccgtgagt ccctctccgg 360

ggtgtgtgaa atcagtggcc gcctctacag actctgctgt cgctgagctt cctagataga 420

aaccaaagca gtgcaagatt cagttcaagg tcctgaaaaa agaaaaacat tttactctgt 480

gtaccttgtg tctttctaaa tttctctctc caaagtaaag ttcaagcatt aaacttagtg 540

tgtttgacct ttttaatttt cttttctttt tcctttttttt tcttttgctt tgttatatgg 600

tggtttgtat ggttcctttt 619

<210> 28
<211> 619
<212> DNA
<213> Homo sapiens
<400> 28
ggatgaagca gaatgaagag taggtaaccc tgaggttgag aggtatattg ttggaccagg 60

gagcaggtaa taaatacatc ctggatagac tcacatgggg aaaaaaacta tgatcttgca 120

tgactaacac atagctagta agatttcttg tcacttacga caaagacatg aattttctcc 180

atcctaacat gactgataca gtgtctctta tttagactat ctcagttagt ctggctgtgc 240

ttgtcctttt tcccacctcc ctcgctgtgc ctgaccctct cttctttcca caggttctca 300

ggcaagagcc acctgctatt gccgaaccgg ccgttgtgct acccgtgagt ccctctccgg 360

ggtgtgtgaa atcagtggcc gcctctacag actctgctgt cgctgagctt cctagataga 420

aaccaaagca gtgcaagatt cagttcaagg tcctgaaaaa agaaaaacat tttactctgt 480

gtaccttgtg tctttctaaa tttctctctc caaaataaag ttcaagcatt aaacttagtg 540

tgtttgacct ttttaatttt cttttctttt tccctttttt tcttttgctt tgttatatgg 600

tggtttgtat ggttcctttt 619

<210> 29
<211> 19
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR primer
<400> 29
ggatgaagca gaatgaaga 19

<210> 30
<211> 19
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR primer
<400> 30
aaaggaacca tacaaacca 19

<210> 31

<211> 18
<212> DNA
<213> Artificial Sequence
<220>
<223> Sequencing primer
<400> 31
gttagtctgg ctgtgctt 18

<210> 32
<211> 1052
<212> DNA
<213> Homo sapiens
<400> 32
gggctactga gtttggtgaa aagataagac tcctgaaaat tctattgatt ctcttttgaa 60

cttctttctt aaattagttt tatgatggac ttggctctca ttggtatttc ccaagattat 120

ggagatggga tagtgatgtc tgacaagtac ctaagatgct aagttgaagg tctaaaattc 180

catcctaaaa gcaaataatt actctatcat ctacgtgccc tttgcttctt aaagttactc 240

aaggaaggca gactaaacag gaaatttact ttggattcaa gaggggcata gagacgctct 300

cagcctgccc atttgccttc atcaacattc ctaaacactg ggcttaaaat gtagtatgag 360

taaactctct cttagtctat ccatctccca ctagcagttt taacatcatc tctagttatt 420

aaccttggct caatggcttt ctcctctttt tttatacaga atttattggc ttgagacgct 480

gtttaatggg tttggggaga tgcagggatc actgcaatgt ggatgaaaaa gagatacaga 540

aatgcaagat gaaaaaatgt tgtgttggac caaaagtggt taaattgatt aaaaactacc 600

tgcaatatgg aacaccaaat gtacttaatg aagacgtcca agaaatgcta aaacctgcca 660

agaattctag tgctgtgata caaagaaaac atattttatc tgttctcccc caaatcaaaa 720

gcactagctt ttttgctaat accaactttg tcatcattcc aaatgccacc cctatgaact 780

ctgccaccat cagcactatg accccaggac agatcacata cactgctact tctaccaaga 840

gtaacaccaa agaaagcaga gattctgcca ctgcctcgcc accaccagca ccacctccac 900

caaacatact gccaacacca tcactggagc tagaggaagc agaagagcag taatgtggat 960

ctttccctta aaactccaag ttcctctcta tttttgctat ctataaaatg acatagaact 1020

gtttcctctg tcatcagtca ttcaataaac ac 1052

<210> 33
<211> 1049
<212> DNA
<213> Homo sapiens
<400> 33
gggctactga gtttggtgaa aagataagac tcctgaaaat tctattgatt ctcttttgaa 60

cttctttctt aaattagttt tatgatggac ttggctctca ttggtatttc ccaagattat 120

ggagatggga tagtgatgtc tgacaagtac ctaagatgct aagttgaagg tctaaaattc 180

catcctaaaa gcaaataatt actctatcat ctacgtgccc tttgcttctt aaagttactc 240

aaggaaggca gactaaacag gaaatttact ttggattcaa gaggggcata gagacgctct 300

cagcctgccc atttgccttc atcaacattc ctaaacactg ggcttaaaat gtagtatgag 360

taaactctct cttagtctat ccatctccca ctagcagttt taacatcatc tctagttatt 420

aaccttggct caatggcttt ctcttttttt atacagaatt tattggcttg agacgctgtt 480

taatgggttt ggggagatgc agggatcact gcaatgtgga tgaaaaagag atacagaaat 540

gcaagatgaa aaaatgttgt gttggaccaa aagtggttaa attgattaaa aactacctgc 600

aatatggaac accaaatgta cttaatgaag acgtccaaga aatgctaaaa cctgccaaga 660

attctagtgc tgtgatacaa agaaaacata ttttatctgt tctcccccaa atcaaaagca 720

ctagcttttt tgctaatacc aactttgtca tcattccaaa tgccacccct atgaactctg 780

ccaccatcag cactatgacc ccaggacaga tcacatacac tgctacttct accaagagta 840

acaccaaaga aagcagagat tctgccactg cctcgccacc accagcacca cctccaccaa 900

acatactgcc aacaccatca ctggagctag aggaagcaga agagcagtaa tgtggatctt 960

tcccttaaaa ctccaagttc ctctctattt ttgctatcta taaaatgaca tagaactgtt 1020

tcctctgtca tcagtcattc aataaacac 1049

<210> 34
<211> 18
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR primer
<400> 34
ggctactgag tttggtga 18

<210> 35
<211> 21
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR primer
<400> 35
gtgtttattg aatgactgat g 21

<210> 36
<211> 18
<212> DNA
<213> Artificial Sequence
<220>
<223> Sequencing primer
<400> 36
caaggaaggc agactaaa 18

<210> 37
<211> 552
<212> DNA
<213> Homo sapiens
<220>
<221> CDS
<222> (1)..(552)
<223> Coding sequence for the variant human DEFB129 gene
<400> 37

```
atg aag ctc ctt ttt cct atc ttt gcc agc ctc atg cta cag tac cag      48
Met Lys Leu Leu Phe Pro Ile Phe Ala Ser Leu Met Leu Gln Tyr Gln
1               5                   10                  15

gtg aac aca gaa ttt att ggc ttg aga cgc tgt tta atg ggt ttg ggg      96
Val Asn Thr Glu Phe Ile Gly Leu Arg Arg Cys Leu Met Gly Leu Gly
            20                  25                  30

aga tgc agg gat cac tgc aat gtg gat gaa aaa gag ata cag aaa tgc     144
Arg Cys Arg Asp His Cys Asn Val Asp Glu Lys Glu Ile Gln Lys Cys
        35                  40                  45

aag atg aaa aaa tgt tgt gtt gga cca aaa gtg gtt aaa ttg att aaa     192
Lys Met Lys Lys Cys Cys Val Gly Pro Lys Val Val Lys Leu Ile Lys
    50                  55                  60

aac tac ctg caa tat gga aca cca aat gta ctt aat gaa gac gtc caa     240
Asn Tyr Leu Gln Tyr Gly Thr Pro Asn Val Leu Asn Glu Asp Val Gln
65                  70                  75                  80

gaa atg cta aaa cct gcc aag aat tct agt gct gtg ata caa aga aaa     288
Glu Met Leu Lys Pro Ala Lys Asn Ser Ser Ala Val Ile Gln Arg Lys
                85                  90                  95

cat att tta tct gtt ctc ccc caa atc aaa agc act agc ttt ttt gct     336
His Ile Leu Ser Val Leu Pro Gln Ile Lys Ser Thr Ser Phe Phe Ala
            100                 105                 110

aat acc aac ttt gtc atc att cca aat gcc acc cct atg aac tct gcc     384
Asn Thr Asn Phe Val Ile Ile Pro Asn Ala Thr Pro Met Asn Ser Ala
        115                 120                 125

acc atc agc act atg acc cca gga cag atc aca tac act gct act tct     432
Thr Ile Ser Thr Met Thr Pro Gly Gln Ile Thr Tyr Thr Ala Thr Ser
    130                 135                 140

acc aag agt aac acc aaa gaa agc aga gat tct gcc act gcc tcg cca     480
Thr Lys Ser Asn Thr Lys Glu Ser Arg Asp Ser Ala Thr Ala Ser Pro
145                 150                 155                 160

cca cca gca cca cct cca cca aac ata ctg cca aca cca tca ctg gag     528
Pro Pro Ala Pro Pro Pro Pro Asn Ile Leu Pro Thr Pro Ser Leu Glu
                165                 170                 175

cta gag gaa gca gaa gag cag taa                                     552
Leu Glu Glu Ala Glu Glu Gln
            180
```

<210> 38
<211> 183

<212> PRT
<213> Homo sapiens
<400> 38

```
Met Lys Leu Leu Phe Pro Ile Phe Ala Ser Leu Met Leu Gln Tyr Gln
1               5                   10                  15

Val Asn Thr Glu Phe Ile Gly Leu Arg Arg Cys Leu Met Gly Leu Gly
            20                  25                  30

Arg Cys Arg Asp His Cys Asn Val Asp Glu Lys Glu Ile Gln Lys Cys
        35                  40                  45

Lys Met Lys Lys Cys Cys Val Gly Pro Lys Val Val Lys Leu Ile Lys
    50                  55                  60

Asn Tyr Leu Gln Tyr Gly Thr Pro Asn Val Leu Asn Glu Asp Val Gln
65                  70                  75                  80

Glu Met Leu Lys Pro Ala Lys Asn Ser Ser Ala Val Ile Gln Arg Lys
                85                  90                  95

His Ile Leu Ser Val Leu Pro Gln Ile Lys Ser Thr Ser Phe Phe Ala
            100                 105                 110

Asn Thr Asn Phe Val Ile Ile Pro Asn Ala Thr Pro Met Asn Ser Ala
        115                 120                 125

Thr Ile Ser Thr Met Thr Pro Gly Gln Ile Thr Tyr Thr Ala Thr Ser
    130                 135                 140

Thr Lys Ser Asn Thr Lys Glu Ser Arg Asp Ser Ala Thr Ala Ser Pro
145                 150                 155                 160

Pro Pro Ala Pro Pro Pro Pro Asn Ile Leu Pro Thr Pro Ser Leu Glu
                165                 170                 175

Leu Glu Glu Ala Glu Glu Gln
            180
```

<210> 39
<211> 552
<212> DNA
<213> Homo sapiens
<220>
<221> CDS
<222> (1)..(552)
<223> Coding sequence for the human DEFB129 gene
<400> 39

atg aag ctc ctt ttt cct atc ttt gcc agc ctc atg cta cag tac cag 48
Met Lys Leu Leu Phe Pro Ile Phe Ala Ser Leu Met Leu Gln Tyr Gln
1 5 10 15

gtg aac aca gaa ttt att ggc ttg aga cgc tgt tta atg ggt ttg ggg 96
Val Asn Thr Glu Phe Ile Gly Leu Arg Arg Cys Leu Met Gly Leu Gly
20 25 30

aga tgc agg gat cac tgc aat gtg gat gaa aaa gag ata cag aaa tgc 144
Arg Cys Arg Asp His Cys Asn Val Asp Glu Lys Glu Ile Gln Lys Cys
35 40 45

aag atg aaa aaa tgt tgt gtt gga cca aaa gtg gtt aaa ttg att aaa 192
Lys Met Lys Lys Cys Cys Val Gly Pro Lys Val Val Lys Leu Ile Lys
50 55 60

aac tac cta caa tat gga aca cca aat gta ctt aat gaa gac gtc caa 240
Asn Tyr Leu Gln Tyr Gly Thr Pro Asn Val Leu Asn Glu Asp Val Gln
65 70 75 80

gaa atg cta aaa cct gcc aag aat tct agt gct gtg ata caa aga aaa 288
Glu Met Leu Lys Pro Ala Lys Asn Ser Ser Ala Val Ile Gln Arg Lys
85 90 95

cat att tta tct gtt ctc ccc caa atc aaa agc act agc ttt ttt gct 336
His Ile Leu Ser Val Leu Pro Gln Ile Lys Ser Thr Ser Phe Phe Ala
100 105 110

aat acc aac ttt gtc atc att cca aat gcc acc cct atg aac tct gcc 384
Asn Thr Asn Phe Val Ile Ile Pro Asn Ala Thr Pro Met Asn Ser Ala
115 120 125

acc atc agc act atg acc cca gga cag atc aca tac act gct act tct 432
Thr Ile Ser Thr Met Thr Pro Gly Gln Ile Thr Tyr Thr Ala Thr Ser
130 135 140

acc aag agt aac acc aaa gaa agc aga gat tct gcc act gcc tcg cca 480
Thr Lys Ser Asn Thr Lys Glu Ser Arg Asp Ser Ala Thr Ala Ser Pro
145 150 155 160

cca cca gca cca cct cca cca aac ata ctg cca aca cca tca ctg gag 528
Pro Pro Ala Pro Pro Pro Pro Asn Ile Leu Pro Thr Pro Ser Leu Glu
165 170 175

cta gag gaa gca gaa gag cag taa 552
Leu Glu Glu Ala Glu Glu Gln
180

<210> 40
<211> 183
<212> PRT
<213> Homo sapiens
<400> 40

Met Lys Leu Leu Phe Pro Ile Phe Ala Ser Leu Met Leu Gln Tyr Gln
1 5 10 15

Val Asn Thr Glu Phe Ile Gly Leu Arg Arg Cys Leu Met Gly Leu Gly
20 25 30

Arg Cys Arg Asp His Cys Asn Val Asp Glu Lys Glu Ile Gln Lys Cys
35 40 45

Lys Met Lys Lys Cys Cys Val Gly Pro Lys Val Val Lys Leu Ile Lys
50 55 60

Asn Tyr Leu Gln Tyr Gly Thr Pro Asn Val Leu Asn Glu Asp Val Gln
65 70 75 80

Glu Met Leu Lys Pro Ala Lys Asn Ser Ser Ala Val Ile Gln Arg Lys
85 90 95

His Ile Leu Ser Val Leu Pro Gln Ile Lys Ser Thr Ser Phe Phe Ala
100 105 110

Asn Thr Asn Phe Val Ile Ile Pro Asn Ala Thr Pro Met Asn Ser Ala
115 120 125

Thr Ile Ser Thr Met Thr Pro Gly Gln Ile Thr Tyr Thr Ala Thr Ser
130 135 140

Thr Lys Ser Asn Thr Lys Glu Ser Arg Asp Ser Ala Thr Ala Ser Pro
145 150 155 160

Pro Pro Ala Pro Pro Pro Pro Asn Ile Leu Pro Thr Pro Ser Leu Glu
165 170 175

Leu Glu Glu Ala Glu Glu Gln
180

<210> 41
<211> 372
<212> DNA
<213> Homo sapiens
<220>
<221> CDS
<222> (1)..(372)
<223> Coding sequence for the variant human DEFB118 gene
<400> 41
atg aaa ctc ctg ctg ctg gct ctt cct atg ctt gtg ctc cta ccc caa 48
Met Lys Leu Leu Leu Leu Ala Leu Pro Met Leu Val Leu Leu Pro Gln
1 5 10 15

gtg atc cca gcc tat agt ggt gaa aaa aaa tgc tgg aac aga tca ggg 96
Val Ile Pro Ala Tyr Ser Gly Glu Lys Lys Cys Trp Asn Arg Ser Gly
20 25 30

cac cgc agg aaa caa tgc aaa gat gga gaa gca gtg aaa gat aca tgc 144
His Arg Arg Lys Gln Cys Lys Asp Gly Glu Ala Val Lys Asp Thr Cys

35 40 45

aaa aat ctt cga gct tgc tgc att cca tcc aat gaa gac cac agg cga 192
Lys Asn Leu Arg Ala Cys Cys Ile Pro Ser Asn Glu Asp His Arg Arg
50 55 60

gtt cct gcg aca tct ccc aca ccc ttg agt gac tca aca cca gga att 240
Val Pro Ala Thr Ser Pro Thr Pro Leu Ser Asp Ser Thr Pro Gly Ile
65 70 75 80

att gat gat att tta aca gta agg ttc acg aca gac tac ttt gaa gta 288
Ile Asp Asp Ile Leu Thr Val Arg Phe Thr Thr Asp Tyr Phe Glu Val
85 90 95

agc agc aag aaa gat atg gtt gaa gag tct gag gcg gga agg gga act 336
Ser Ser Lys Lys Asp Met Val Glu Glu Ser Glu Ala Gly Arg Gly Thr
100 105 110

gag acc tct ctt cca aat gtt cac cat agc tca tga 372
Glu Thr Ser Leu Pro Asn Val His His Ser Ser
115 120

<210> 42
<211> 123
<212> PRT
<213> Homo sapiens
<400> 42

Met Lys Leu Leu Leu Leu Ala Leu Pro Met Leu Val Leu Leu Pro Gln
1 5 10 15

Val Ile Pro Ala Tyr Ser Gly Glu Lys Lys Cys Trp Asn Arg Ser Gly
20 25 30

His Arg Arg Lys Gln Cys Lys Asp Gly Glu Ala Val Lys Asp Thr Cys
35 40 45

Lys Asn Leu Arg Ala Cys Cys Ile Pro Ser Asn Glu Asp His Arg Arg
50 55 60

Val Pro Ala Thr Ser Pro Thr Pro Leu Ser Asp Ser Thr Pro Gly Ile
65 70 75 80

Ile Asp Asp Ile Leu Thr Val Arg Phe Thr Thr Asp Tyr Phe Glu Val
85 90 95

Ser Ser Lys Lys Asp Met Val Glu Glu Ser Glu Ala Gly Arg Gly Thr
100 105 110

Glu Thr Ser Leu Pro Asn Val His His Ser Ser
115 120

```
<210>  43
<211>  372
<212>  DNA
<213>  Homo sapiens
<220>
<221>  CDS
<222>  (1)..(372)
<223>  Coding sequence of the human DEFB118 gene
<400>  43
atg aaa ctc ctg ctg ctg gct ctt cct atg ctt gtg ctc cta ccc caa      48
Met Lys Leu Leu Leu Leu Ala Leu Pro Met Leu Val Leu Leu Pro Gln
1               5                   10                  15

gtg atc cca gcc tat agt ggt gaa aaa aaa tgc tgg aac aga tca ggg      96
Val Ile Pro Ala Tyr Ser Gly Glu Lys Lys Cys Trp Asn Arg Ser Gly
            20                  25                  30

cac tgc agg aaa caa tgc aaa gat gga gaa gca gtg aaa gat aca tgc     144
His Cys Arg Lys Gln Cys Lys Asp Gly Glu Ala Val Lys Asp Thr Cys
        35                  40                  45

aaa aat ctt cga gct tgc tgc att cca tcc aat gaa gac cac agg cga     192
Lys Asn Leu Arg Ala Cys Cys Ile Pro Ser Asn Glu Asp His Arg Arg
    50                  55                  60

gtt cct gcg aca tct ccc aca ccc ttg agt gac tca aca cca gga att     240
Val Pro Ala Thr Ser Pro Thr Pro Leu Ser Asp Ser Thr Pro Gly Ile
65                  70                  75                  80

att gat gat att tta aca gta agg ttc acg aca gac tac ttt gaa gta     288
Ile Asp Asp Ile Leu Thr Val Arg Phe Thr Thr Asp Tyr Phe Glu Val
                85                  90                  95

agc agc aag aaa gat atg gtt gaa gag tct gag gcg gga agg gga act     336
Ser Ser Lys Lys Asp Met Val Glu Glu Ser Glu Ala Gly Arg Gly Thr
            100                 105                 110

gag acc tct ctt cca aat gtt cac cat agc tca tga                     372
Glu Thr Ser Leu Pro Asn Val His His Ser Ser
        115                 120


<210>  44
<211>  123
<212>  PRT
<213>  Homo sapiens
<400>  44

Met Lys Leu Leu Leu Leu Ala Leu Pro Met Leu Val Leu Leu Pro Gln
1               5                   10                  15


Val Ile Pro Ala Tyr Ser Gly Glu Lys Lys Cys Trp Asn Arg Ser Gly
            20                  25                  30


His Cys Arg Lys Gln Cys Lys Asp Gly Glu Ala Val Lys Asp Thr Cys
        35                  40                  45


Lys Asn Leu Arg Ala Cys Cys Ile Pro Ser Asn Glu Asp His Arg Arg
```

50 55 60

Val Pro Ala Thr Ser Pro Thr Pro Leu Ser Asp Ser Thr Pro Gly Ile
65 70 75 80

Ile Asp Asp Ile Leu Thr Val Arg Phe Thr Thr Asp Tyr Phe Glu Val
85 90 95

Ser Ser Lys Lys Asp Met Val Glu Glu Ser Glu Ala Gly Arg Gly Thr
100 105 110

Glu Thr Ser Leu Pro Asn Val His His Ser Ser
115 120

<210> 45
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR primer
<400> 45
aggttgagta tttgccagac 20

<210> 46
<211> 19
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR primer
<400> 46
aggacagggg tgagtgata 19

<210> 47
<211> 246
<212> DNA
<213> Homo sapiens
<220>
<221> CDS
<222> (1)..(246)
<223> Coding sequence for the variant human DEFB126 gene
<400> 47
atg aag tcc cta ctg ttc acc ctt gca gtt ttt atg ctc ctg gcc caa 48
Met Lys Ser Leu Leu Phe Thr Leu Ala Val Phe Met Leu Leu Ala Gln
1 5 10 15

ttg gtc tca ggt aat tgg tat gtg aaa aag tgt cta aac gac gtt gga 96
Leu Val Ser Gly Asn Trp Tyr Val Lys Lys Cys Leu Asn Asp Val Gly
20 25 30

att tgc aag aag aag tgc aaa cct gaa gag atg cat gta aag aat ggt 144
Ile Cys Lys Lys Lys Cys Lys Pro Glu Glu Met His Val Lys Asn Gly
35 40 45

tgg gca atg tgc ggc aaa ggg act gct gtg ttc cag ctg aca gac gtg 192

Trp Ala Met Cys Gly Lys Gly Thr Ala Val Phe Gln Leu Thr Asp Val
50 55 60

cta att atc ctg ttt tct gtg tcc aga caa aga cta caa gaa ttt caa 240
Leu Ile Ile Leu Phe Ser Val Ser Arg Gln Arg Leu Gln Glu Phe Gln
65 70 75 80

cag taa 246
Gln

<210> 48
<211> 81
<212> PRT
<213> Homo sapiens
<400> 48

Met Lys Ser Leu Leu Phe Thr Leu Ala Val Phe Met Leu Leu Ala Gln
1 5 10 15

Leu Val Ser Gly Asn Trp Tyr Val Lys Lys Cys Leu Asn Asp Val Gly
20 25 30

Ile Cys Lys Lys Lys Cys Lys Pro Glu Glu Met His Val Lys Asn Gly
35 40 45

Trp Ala Met Cys Gly Lys Gly Thr Ala Val Phe Gln Leu Thr Asp Val
50 55 60

Leu Ile Ile Leu Phe Ser Val Ser Arg Gln Arg Leu Gln Glu Phe Gln
65 70 75 80

Gln

<210> 49
<211> 336
<212> DNA
<213> Homo sapiens
<220>
<221> CDS
<222> (1)..(336)
<223> Coding sequence of the human DEFB126 gene
<400> 49
atg aag tcc cta ctg ttc acc ctt gca gtt ttt atg ctc ctg gcc caa 48
Met Lys Ser Leu Leu Phe Thr Leu Ala Val Phe Met Leu Leu Ala Gln
1 5 10 15

ttg gtc tca ggt aat tgg tat gtg aaa aag tgt cta aac gac gtt gga 96
Leu Val Ser Gly Asn Trp Tyr Val Lys Lys Cys Leu Asn Asp Val Gly
20 25 30

att tgc aag aag aag tgc aaa cct gaa gag atg cat gta aag aat ggt 144
Ile Cys Lys Lys Lys Cys Lys Pro Glu Glu Met His Val Lys Asn Gly

```
        35                  40                  45

tgg gca atg tgc ggc aaa caa agg gac tgc tgt gtt cca gct gac aga      192
Trp Ala Met Cys Gly Lys Gln Arg Asp Cys Cys Val Pro Ala Asp Arg
    50                  55                  60

cgt gct aat tat cct gtt ttc tgt gtc cag aca aag act aca aga att      240
Arg Ala Asn Tyr Pro Val Phe Cys Val Gln Thr Lys Thr Thr Arg Ile
65                  70                  75                  80

tca aca gta aca gca aca aca gca aca aca act ttg atg atg act act      288
Ser Thr Val Thr Ala Thr Thr Ala Thr Thr Thr Leu Met Met Thr Thr
                85                  90                  95

gct tcg atg tct tcg atg gct cct acc ccc gtt tct ccc act ggt tga      336
Ala Ser Met Ser Ser Met Ala Pro Thr Pro Val Ser Pro Thr Gly
            100                 105                 110
```

```
<210>  50
<211>  111
<212>  PRT
<213>  Homo sapiens
<400>  50

Met Lys Ser Leu Leu Phe Thr Leu Ala Val Phe Met Leu Leu Ala Gln
1               5                   10                  15

Leu Val Ser Gly Asn Trp Tyr Val Lys Lys Cys Leu Asn Asp Val Gly
            20                  25                  30

Ile Cys Lys Lys Lys Cys Lys Pro Glu Glu Met His Val Lys Asn Gly
        35                  40                  45

Trp Ala Met Cys Gly Lys Gln Arg Asp Cys Cys Val Pro Ala Asp Arg
    50                  55                  60

Arg Ala Asn Tyr Pro Val Phe Cys Val Gln Thr Lys Thr Thr Arg Ile
65                  70                  75                  80

Ser Thr Val Thr Ala Thr Thr Ala Thr Thr Thr Leu Met Met Thr Thr
                85                  90                  95

Ala Ser Met Ser Ser Met Ala Pro Thr Pro Val Ser Pro Thr Gly
            100                 105                 110
```

```
<210>  51
<211>  20
<212>  DNA
<213>  Artificial Sequence
<220>
<223>  PCR primer
<400>  51
aatggtgaga aagatgacag                                                20
```

<210> 52
<211> 18
<212> DNA
<213> Artificial Sequence
<220>
<223> PCR primer
<400> 52
gttgaatgga gggaaagt 18

<210> 53
<211> 18
<212> DNA
<213> Artificial Sequence
<220>
<223> Sequencing primer
<400> 53
gtaggtattt atgattag 18

<210> 54
<211> 334
<212> DNA
<213> Homo sapiens
<220>
<221> CDS
<222> (1)..(333)
<223> Coding sequence for the variant human DEFB126 gene
<400> 54
atg aag tcc cta ctg ttc acc ctt gca gtt ttt atg ctc ctg gcc caa 48
Met Lys Ser Leu Leu Phe Thr Leu Ala Val Phe Met Leu Leu Ala Gln
1 5 10 15

ttg gtc tca ggt aat tgg tat gtg aaa aag tgt cta aac gac gtt gga 96
Leu Val Ser Gly Asn Trp Tyr Val Lys Lys Cys Leu Asn Asp Val Gly
20 25 30

att tgc aag aag aag tgc aaa cct gaa gag atg cat gta aag aat ggt 144
Ile Cys Lys Lys Lys Cys Lys Pro Glu Glu Met His Val Lys Asn Gly
35 40 45

tgg gca atg tgc ggc aaa caa agg gac tgc tgt gtt cca gct gac aga 192
Trp Ala Met Cys Gly Lys Gln Arg Asp Cys Cys Val Pro Ala Asp Arg
50 55 60

cgt gct aat tat cct gtt ttc tgt gtc cag aca aag act aca aga att 240
Arg Ala Asn Tyr Pro Val Phe Cys Val Gln Thr Lys Thr Thr Arg Ile
65 70 75 80

tca aca gta aca gca aca aca gca aca aca act ttg atg atg act act 288
Ser Thr Val Thr Ala Thr Thr Ala Thr Thr Thr Leu Met Met Thr Thr
85 90 95

gct tcg atg tct tcg atg gct cct acc cgt ttc tcc cac tgg ttg a 334
Ala Ser Met Ser Ser Met Ala Pro Thr Arg Phe Ser His Trp Leu
100 105 110

<210> 55

<211> 111
<212> PRT
<213> Homo sapiens
<400> 55

Met Lys Ser Leu Leu Phe Thr Leu Ala Val Phe Met Leu Leu Ala Gln
1 5 10 15

Leu Val Ser Gly Asn Trp Tyr Val Lys Lys Cys Leu Asn Asp Val Gly
20 25 30

Ile Cys Lys Lys Lys Cys Lys Pro Glu Glu Met His Val Lys Asn Gly
35 40 45

Trp Ala Met Cys Gly Lys Gln Arg Asp Cys Cys Val Pro Ala Asp Arg
50 55 60

Arg Ala Asn Tyr Pro Val Phe Cys Val Gln Thr Lys Thr Thr Arg Ile
65 70 75 80

Ser Thr Val Thr Ala Thr Thr Ala Thr Thr Thr Leu Met Met Thr Thr
85 90 95

Ala Ser Met Ser Ser Met Ala Pro Thr Arg Phe Ser His Trp Leu
100 105 110

<210> 56
<211> 50
<212> DNA
<213> Artificial Sequence
<220>
<223> snapshot primer
<400> 56
tttttttttt tttttttttt tttttttttt tttgctcaat ggctttctct 50